VASILE TAMAŞ (coord.)
DAN BRÂNZEI
CONSTANTIN SMADICI
JULIEN MOSCOVICI
RODICA BRÂNZEI

MATEMATICI GENERALE PENTRU ECONOMIŞTI

87.837

BCU IASI / CENTRAL UNIVERSITY LIBRARY

CUPRINS

INTRODUCERE 7

CAP.I MULŢIMI, RELAŢII, FUNCŢII

§ 1. Mulţimi 9
§ 2. Algebre de mulţimi 13
§ 3. Relaţii,funcţii 16
§ 4. Elemente de analiză combinatorie 25
Întrebări de autocontrol, exerciţii 29

CAP.II ELEMENTE DE ALGEBRĂ

§ 1. Spaţii vectoriale (liniare) 30
§ 2. Subspaţii liniare 33
§ 3. Dependenţă liniară, baze,coordonate 35
§ 4. Operatori liniari 44
§ 5. Forme liniare, biliniare şi pătratice 47
§ 6. Spaţii euclidiene 51
§ 7. Sisteme de ecuaţii liniare 53
§ 8. Polinoame de interpolare 60
Întrebări de autocontrol, exerciţii 63

CAP.III SERII

§ 1. Serii numerice 66
§ 2. Serii de funcţii 78
Întrebări de autocontrol, exerciţii 86

BCU CLUJ "LUCIAN BLAGA" UNIVERSITY LIBRARY

CAP.IV FUNCŢII DE MAI MULTE VARIABILE

§ 1. Generalităţi 87
§ 2. Derivate parţiale. Extreme 90
§ 3. Metoda celor mai mici pătrate 95
§ 4. Extreme legate 97
Întrebări de autocontrol, exerciţii 100

CAP.V EXTINDEREA NOŢIUNII DE INTEGRALĂ

§ 1. Intregala Riemann-Stieltjes 102
§ 2. Integrale improprii 104
§ 3. Funcţiile beta şi gama, integrala Euler-Poisson 108
§ 4. Integrala dublă 111
§ 5. Formule de integrare aproximativă 116
Întrebări de autocontrol, exerciţii 124

CAP.VI ECUAŢII DIFERENŢIALE

§ 1. Generalităţi 126
§ 2. Ecuaţii diferenţiale ordinare de ordinul întîi. Generalităţi 129
§ 3. Tipuri speciale de ecuaţii diferenţiale ordinare de ordinul întîi . . . 132
§ 4. Ecuaţii diferenţiale ordinare de ordin superior 145
Întrebări de autocontrol, exerciţii 161

CAP.VII ELEMENTE DE CALCULUL PROBABILITĂŢILOR

§ 1. Introducere 162
§ 2. Evenimente. Operaţii cu evenimente 164
§ 3. Cîmp de evenimente 169
§ 4. Probabilitate. Cîmp de probabilitate 175
§ 5. Variabile aleatoare 214
§ 6. Caracteristici numerice ale unei variabile aleatoare 248
§ 7. Repartiţii probabilstice discrete vizuale 272
§ 8. Repartiţii uzuale de tip continuu 284
§ 9. Legea numerelor mari 319

BIBLIOGRAFIE 335

BCU CENTRAL UNIVERSITY LIBRARY

INTRODUCERE

Ne aflăm la confluenţa a două milenii, a două ere caracterizate prin tipuri diferite de civilizaţie şi organizare socială. Acest cadru face ca omul să trăiască şi să acţioneze într-un mediu economic extrem de complex aflat într-un proces rapid de transformare profundă şi multiplă.

Cunoscutul matematician român O.Onicescu în lucrarea „Incertitudine şi modelare economică" încă cu mulţi ani în urmă sublinia „importanţa competenţei şi lucidităţii, singurele care pot aduce bunele intenţii şi entuziasmul din cîmpul visurilor în cel al practicii".

Nu mai constituie o noutate faptul că activitatea modernă în domeniul producţiei industriale şi conducerii economice a devenit atît de complexă încît rezolvările tradiţionale nu mai sînt suficiente.

Cea mai mare parte a problemelor întîlnite sînt probleme de optimizare. Pe ansamblu, cele mai multe sînt abordabile şi se pot soluţiona folosind modelarea matematică. Metodele matematice stau la baza gestiunii şi a conducerii ştiinţifice a întreprinderilor, la studiul balanţelor economiilor naţionale, la elaborarea prognozelor privind economia pe termen lung. În toate domeniile de natură economică, financiară, comercială şi de organizare a producţiei, metodele matematice şi-au arătat deplina eficienţă în numeroase ţări avansate din lume. Importanţa metodelor ştiinţifice, în general, şi a celor matematice, în particular, dorim să o evidenţiem prin cîteva consideraţii legate de noţiunea de informaţie şi rolul ei în viaţa modernă.

În orice acţiune iniţiată de un întreprinzător în orice moment poate să apară aşa-numitul risc. Este util să încercăm să descifrăm mecanismele apariţiei şi declanşării riscului în lumea afacerilor. Vom fi obligaţi să recunoaştem importanţa probabilităţii producerii unui eveniment nefavorabil proiectului întreprins. Deci sîntem în situaţia de a acţiona între certitudine şi hazard, între care există nenumărate trepte.

Complexitatea crescîndă a fenomenelor social-economice precum şi derularea lor extraordinar de rapidă, i-au făcut pe unii economişti să numească perioada prin care treceau „era incertitudinii". Pentru a putea preveni sau contracara efectele neprevăzutului, avem nevoie de „informaţie". Întreaga societate de azi se caracterizează printr-o tranziţie evidentă de la societatea industrială la societatea informaţională. Trecerea la economia de piaţă impune cu o putere şi mai mare necesitatea obţinerii şi prelucrării informaţiei în vederea folosirii ei viitoare. Acest lucru nu este posibil fără folosirea intensivă a modelării matematice.

Acesta este motivul pentru care un viitor economist va avea mare nevoie de cunoştinţe de matematică. Volumul de faţă îşi propune să completeze materia de liceu cu elementele strict necesare activităţii de documentare şi cercetare efectuate de economişti sau viitori economişti.

În această primă parte am inclus aparatul matematic necesar în procesul modelării, sperînd că într-un volum următor să prezentăm şi tipuri de probleme modelate însoţite de metodele adecvate de soluţionare.

Pentru a veni în sprijinul celor interesaţi la sfîrşit de capitol sau chiar la sfîrşit de paragraf (cînd a fost cazul) am propus cîteva întrebări de autocontrol şi exerciţii numerice spre rezolvare.

Pentru clarificarea terminologiei şi pentru introducerea unor noutăţi, primul capitol a fost dedicat teoriei mulţimilor. În capitolul al II-lea am prezentat cu atenţie (dar pe scurt) elementele din teoria spaţiilor liniare, operatorilor liniari, formelor liniare, biliniare şi pătratice, sistemelor de ecuaţii. Am considerat nimerit să tratăm şi problema interpolării. Am introdus un scurt capitol de serii, unul de funcţii de mai multe variabile, precum şi unul consacrat extinderii noţiunii de integrală: *Natura problemelor studiate de un economist cu cunoştinţe legate de ecuaţii diferenţiale.* Din acest motiv capitolul VI este rezervat ecuaţiilor diferenţiale ordinare, limitîndu-ne la elementele strict necesare şi accesibile. În fine, în capitolul VII, care este cel mai întins, prezentăm teoria probabilităţilor, construindu-ne baza teoretică necesară multor capitole viitoare.

La realizarea lucrării, contribuţia autorilor este următoarea:

Cap. I	V.Tamaş şi R.Brânzei
Cap. II	V.Tamaş şi D.Brânzei
Cap. III	V.Tamaş şi D.Brânzei
Cap. IV	V.Tamaş
Cap. V	V.Tamaş şi D.Brânzei
Cap. VI	J.Moscovici şi V.Tamaş
Cap. VII	C.Smadici

Asamblarea finală, armonizarea notaţiilor, precum şi realizarea unităţii materialului a fost urmărită de V.Tamaş.

B.C.U. „M. EMINESCU" IAŞI

CAPITOLUL I

MULŢIMI, RELAŢII, FUNCŢII

§ 1. MULŢIMI

Noţiunea de mulţime reprezintă matematizarea noţiunii intuitive de colecţie sau grămadă. Izvoarele teoriei mulţimilor se trag din studiile făcute asupra numerelor sau punctelor unui segment de dreaptă. Conceptul abstract de mulţime a fost abordat de mulţi matematicieni renumiţi din a doua jumătate a secolului XIX. Printre primii care au adus o contribuţie decisivă în această direcţie a fost Georg Cantor (1845-1918) care în perioada 1870-1880 a scris primele lucrări dedicate acestei noţiuni. Acum la un secol de la începuturi noţiunea de mulţime stă la baza întregului edificiu al matematicii şi deci şi la baza tuturor aplicaţiilor sale. Abstractizarea noţiunii fizice de colecţie l-a condus pe Cantor la afirmaţia că „se numeşte mulţime orice colecţie de obiecte determinate considerată în întregul ei". Prin urmare se înţelege că chiar de la începuturile apariţiei teoriei mulţimilor noţiunea de mulţime a fost considerată distinctă de cea de colecţie. Urmîndu-se această cale s-a constituit aşa numita teorie naivă sau teorie intuitivă a mulţimilor care cu toate progresele făcute s-a dovedit a fi marcată de numeroase contradicţii logice. Recunoaşterea unor asemenea lipsuri a teoriei intuitive nu înseamnă că vom renunţa la ea şi o vom folosi permanent cu precauţiile care se impun.

Pentru noi noţiunea de mulţime este o noţiune primară care nu se defineşte ci numai se descrie prin a spune că mulţimea este o existenţă matematică formată din elemente despre care se spune că aparţin mulţimii. Elementele care formează o mulţime trebuiesc gîndite ca fiind bine determinate, fiecare avînd individualitatea

sa şi nu există nici o ambiguitate în ceea ce priveşte criteriile care ne arată dacă un element aparţine sau nu unei mulţimi.

Aşa dar o mulţime este determinată dacă se cunosc toate elementele sale sau se cunoaşte ce proprietate au toate elementele sale. Un element al unei mulţimi se consideră o singură dată, deci elementele unei mulţimi sînt distincte.

Vom nota mulţimile cu litere mari A, B, C, ..., Z iar elementele unei mulţimi le vom nota cu litere mici cu sau fără indici.

O mulţime poate fi dată fie indicînd fiecare din elementele sale fie specificînd o proprietate pe care o au elementele sale. Deci:

$A = \{a, b, c, \ldots\}$ sau

$A = \{a_1, a_2, a_3, \ldots, a_n\}$ sau

$A = \{x \mid \mathcal{P}(x)\}$

unde $\mathcal{P}$ este proprietatea comună a elementelor. În prima variantă se spune că mulţimea este dată în extensie iar în a doua în comprehensiune.

Faptul că o mulţime este determinată de elementele sale implică evident că două mulţimi sînt egale dacă şi numai dacă sînt formate din aceleaşi elemente.

Mulţimea care nu conţine nici un element se numeşte mulţime vidă şi se notează cu Φ. Dacă a este un element al mulţimii A se scrie $a \in A$ iar dacă a nu face parte din A se scrie $a \notin A$. Simbolul „$\in$" este simbolul de apartenenţă.

Definiţia 1.1.

Dacă A, B, sînt două mulţimi, vom spune că A este submulţime a lui B dacă oricare ar fi $x \in A$ urmează că $x \in B$.

Se mai spune că A este inclusă în B sau că A este o parte a lui B şi se scrie $A \subset B$. Această relaţie (legătură) între mulţimi se numeşte relaţie de incluziune şi are proprietăţile:

1. $A \subseteq A$ - (reflexivitate)
2. $A \subseteq B$ şi $B \subseteq C \rightarrow A \subseteq C$ - (tranzitivitate)
3. Dacă $A \subseteq B$ şi $B \subseteq A \rightarrow A = B$ (antisimetrie)

Dacă $A \neq B$ şi ŏ $A \subseteq B$ vom spune că A este strict inclusă în B şi vom scrie $A \subset B$.

Conform celor precedente Φ şi A sînt submulţimi ale mulţimii A numite submulţimi improprii.

Submulţimile lui A diferite de Φ şi A se numesc submulţimi proprii

ale lui A. Dacă S este submulţime proprie a mulţimii A avem $S \subset A$.

Mulţimile ale căror elemente sînt la rîndul lor mulţimi se numesc familii sau sisteme. Un sistem foarte important este mulţimea submulţimilor unei mulţimi date T., aceasta se numeşte mulţimea părţilor mulţimii T şi se notează cu $\mathcal{P}(T)$. Deci:

$$P(T) = \{ M \mid M \subseteq T \}$$

În teoria mulţimilor din una, două sau mai multe mulţimi se formează noi mulţimi după anumite reguli pe care le numim operaţii.

Definiţia 1.2.

Fie A şi B două mulţimi. Cu ajutorul lor formăm mulţimile:

$$A \cup B = \{ x \mid x \in A \text{ sau } x \in B \}$$

$$A \cap B = \{ x \mid x \in A \text{ şi } x \in B \}$$

$$A - B = \{ x \mid x \in A \text{ şi } x \notin B \}$$

Prima mulţime se numeşte reuniune, a doua intersecţie iar a treia diferenţa mulţimilor A şi B.

Dintre principalele proprietăţi ale reuniunii şi intersecţiei amintim:

1. Idempotenţa $\quad A \cup A = A \ , \ A \cap A = A$
2. Comutativitatea $\quad A \cup B = B \cup A \ , \ A \cap B = B \cap A$
3. Asociativitatea $\quad \begin{cases} A \cup (B \cup C) = (A \cup B) \cup C \\ A \cap (B \cap C) = (A \cap B) \cap C \end{cases}$
4. Absorbţia $\quad A \cup (A \cap B) = A \ , \ A \cap (A \cup B) = A$
5. Distributivitatea $\quad A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$$

Deoarece proprietăţile de la intersecţie se obţin din cele de la reuniune prin schimbarea semnelor $\cup$ cu $\cap$ şi invers, am căpătat denumirea de formule duale (proprietăţi duale). Dacă $A \cap B = \Phi$ mulţimile A, B se numesc disjuncte.

BUJA CENTRAL UNIVERSITY LIBRARY

Definiţia 1.3.

Fie M o mulţime şi A o submulţime a sa. Se numeşte complementara mulţimii A în raport cu M, mulţimea $C_M A = \{x \in M \mid x \notin A\}$. Dacă admitem că vom avea în vedere numai elementele din $\mathcal{P}(M)$ putem nota $C_M A = \overline{A}$.

Au loc formulele:

$$\overline{A \cup B} = \overline{A} \cap \overline{B}$$

$$\overline{A \cap B} = \overline{A} \cup \overline{B}$$

numite formulele lui de Morgan, şi $\overline{A} = M - A$.

Definiţia 1.4.

Fie A şi B două mulţimi şi să notăm $(a,b) = \{\{a\}, \{a,b\}\}$ numită pereche ordonată de elemnte $a \in A$ şi $b \in B$. Se numeşte produsul cartezian al mulţimilor A şi B şi se notează cu $A \times B$ mulţimea: $A \times B = \{(a,b) \mid a \in A \text{ şi } b \in B\}$. Datorită ordinii fixate în (a,b) este clar că $A \times B \neq B \times A$ şi $A \times B = B \times A$ atunci şi numai atunci cînd $B = A$. Evident că dacă $A_1 \subseteq A$ şi $B_1 \subseteq B$ rezultă $A_1 \times B_1 \subseteq A \times B$. Trecerea de la cazul a două mulţimi la mai multe se poate face folosind inducţia matematică, proprietăţile operaţiilor rămînînd valabile.

Aşa cum deja am observat elementele unei mulţimi pot fi ele însele mulţimi. În astfel de cazuri interesează pe lîngă operaţiile între mulţimi, în general, şi operaţiile între anumite submulţimi ale lor care formează familii de mulţimi.

Pentru multe probleme din matematici faptul că rezultatul operaţiilor cu elemente dintr-o familie face parte tot din familia respectivă este esenţial, de exemplu în teoria probabilităţilor, algebră, analiză matematică etc. Din acest motiv ne oprim puţin asupra lor.

Fie A o mulţime şi $A_1, A_2, \ldots, A_n$ o familie de submulţimi din A, adică o familie de elemente din $\mathcal{P}(A)$. Dacă $A_i, i = \overline{1,n}$ nu sînt toate vide şi $A = \bigcup_{k=1}^{n} A_k$ spunem că familia respectivă formează o *acoperire* finită pentru mulţimea A. Evident că o mulţime poate admite mau multe acoperiri.

Dacă în plus mulţimile A_i sînt disjuncte două cîte două familia $A_1, A_2, \ldots, A_n$ constituie o descompunere finită a mulţimii A sau o *partiţie*.

Noţiunile de acoperire şi partiţie se extind şi în cazurile în care familiile

de mulţimi nu au neapărat un număr finit de elemente. Existenţa unor acoperiri cu elemente din familii cu anumite proprietăţi nu este totdeauna asigurată.

Deoarece partiţiile le vom utiliza şi în cazul finit şi în cazul infinit le vom defini în general.

Definiţia 1.5.

Familia $F = \{A_1, A_2, \ldots\}$ formează o partiţie a mulţimii A dacă:

1. $A_i \neq \Phi$,
2. $A_i \cap A_j = \Phi$, $i \neq j$,
3. $A = \bigcup_{A_k \in F} A_k$.

În matematica modernă noţiunea de partiţie este esenţială deoarece ea reflectă foarte bine operaţia obişnuită de clasificare fără de care nimic nu este posibil în nici un compartiment al activităţii umane.

§ 2. ALGEBRE DE MULŢIMI

Folosind operaţiile de reuniune, intersecţie, diferenţă şi complementariere mai putem defini una nouă numită diferenţă simetrică, notată prin simbolul Δ, a două submulţimi din M.

$$A \Delta B = (A - B) \cup (B - A)$$

Această nouă operaţie are, printre altele, proprietăţile:

1. $A \Delta A = \Phi$
2. $A \Delta B = B \Delta A$
3. $(A \Delta B) \Delta C = A \Delta (B \Delta C)$
4. $A \Delta \Phi = \Phi \Delta A = A$

Consideraţiile următoare se datoresc necesităţii explicării unor termeni folosiţi foarte des fără a fi explicaţi suficient.

Definiţia 1.6.

O familie nevidă I de mulţimi se numeşte inel de mulţimi dacă este închisă în raport cu diferenţa simetrică şi intersecţia unui număr finit de

elemente din I. Prin urmare: $A, B \in I \rightarrow A \Delta B \in I$ şi $A \cap B \in I$.

Observaţia 1. Deoarece $A \cup B = (A \Delta B) \Delta (A \cap B) \rightarrow A \cup B \in I$. Apoi avem $A - B = (A - B) \cup \Phi = (A - A \cap B) \cup (A \cap B - A) = A \Delta (A \cap B)$ deci $A - B \in I$.

Drept consecinţă $A - A \in I$ deci $\Phi \in I$. Putem deci concluziona că orice inel de mulţimi este închis şi la reuniune şi diferenţă şi conţine totdeauna mulţimea vidă.

Această constatare permite să arătăm fără dificultate că inelul de mulţimi poate fi definit şi ca o familie de mulţimi închisă la reuniune şi diferenţă avînd în vedere că:

$$A \Delta B = (A - B) \cup (B - A)$$

$$A \cap B = (A \cup B) - (A \Delta B)$$

Observaţia 2. Dacă notăm cu $A + B = A \Delta B$ şi $AB = A \cap B$ se constată imediat că pentru orice $A, B, C \in I$ au loc:

1. $(A + B) + C = A + (B + C)$
2. $A + \Phi = A$
3. $A + A = \Phi$
4. $A + B = B + A$
5. $(AB)C = A(BC)$
6. $A(B + C) = AB + AC$
7. $AB = BA$
8. $A \cdot A = A$

Aceste proprietăţi arată că I este de fapt un inel algebric obişnuit pe care l-am studiat la liceu.

Dacă inelul I conţine şi o mulţime E cu proprietatea că oricare ar fi $A \in I$ urmează $AE = A$ inelul de mulţimi se numeşte *inel unitar, algebră de mulţimi* sau *cîmp*.

Dacă I este o algebră de mulţimi atunci conform definiţiei urmează că $A \cap E = A$. Cum aceasta trebuie să se întîmple pentru orice A urmează că $A \subset E$.

Aşadar o algebră de mulţimi conţine o mulţime E care le conţine pe toate, este cea mai cuprinzătoare.

Dacă I este un inel de mulţimi oarecare şi dacă nu are unitate prin considerarea reuniunii tuturor mulţimilor din I aceasta devine algebră de mulţimi. Deci dată familia I de mulţimi organizată ca inel de mulţimi o putem organiza ca algebră de mulţimi adăugîndu-i reuniunea. Problema s-ar putea pune şi invers. Dată o mulţime A să-i ataşăm o algebră. Acest lucru este posibil luînd familii F din $\mathcal{P}(A)$ ce pot fi organizate ca inel şi din care să facă parte şi A. Există, deci, mai multe astfel de posibilităţi. Deci este justificată o expresie de forma „fie o mulţime A şi F o algebră de părţi ale sale".

Dacă I este o algebră de părţi a mulţimii A şi $M \in I$ urmează că şi $A - M \in I$ adică complementara oricărei mulţimi M din algebră face parte tot din algebră. Denumirea de algebră este justificată prin faptul că familia respectivă, în raport cu operaţiile: diferenţă simetrică, intersecţia şi complementariere, satisface condiţiile de definiţie a unei algebre.

Definiţia 1.7.

Familia nevidă K de părţi din mulţimea A se numeşte *corp de părţi* dacă este închisă la reuniune finită (număr finit de mulţimi) şi complementariere. Putem spune deci că:

1. $A_1 \in K$ implică $\overline{A_1} \in K$
2. Dacă $A_1, A_2 \in K$ rezultă $A_1 \cup A_2 \in K$

Din condiţiile impuse rezultă imediat că $\Phi \in K$ şi $A \in K$. În adevăr din K nevidă urmează că există $A_1 \in K$ dar atunci conform primei condiţii $A_1 \in K$ şi după cea de a doua $A_1 \cup \overline{A_1}$, care este tocmai A, face parte tot din K.

De asemenea dacă $A_1, A_2 \in K$ atunci $A_1 \cap A_2 \in K$ şi $A_1 - A_2 \in K$. Afirmaţia se justifică avînd în vedere că $A_1 \cap A_2 = \overline{\overline{A_1} \cup \overline{A_2}}$ şi $A_1 - A_2 = A_1 \cap \overline{A_2}$. Putem spune deci că un corp de părţi este o algebră de mulţimi.

Trecerea de la reuniune finită la reuniune numărabilă ne duce la noţiuni corespunzătoare noi.

Definiţia 1.8.

Dacă o familie nevidă F de părţi este un inel închis faţă de reuniunea numărabilă atunci se numeşte σ-inel. Un σ-inel de părţi dintr-o mulţime A

care o conţine se numeşte σ-algebră şi o familie de părţi închisă faţă de reuniunea numărabilă şi complementariere poartă numele de σ-corp.

Deoarece în toate aceste structuri pentru $M \in F$ urmează $MM = M$ şi $M + M = \Phi$ acestea sînt numite reuniuni booleene.

Rolul structurilor introduse mai înainte va fi de mare utiliate în înţelegerea corectă a noţiunilor din capitolul dedicat teoriei probabilităţilor.

§ 3. RELAŢII, FUNCŢII

În matematică, la fel ca şi în natura înconjurătoare, obiectele se găsesc în diferite legături sau relaţii. O situaţie atît de universală, ca cea de relaţie, nu putea rămîne în afara preocupărilor matematicii. Pentru a putea fi folosită ea trebuie formalizată aşa cum am trecut de la ideea de colecţie la cea de mulţime.

Relaţia sau corespondenţa este elementul principal al tuturor consideraţiilor următoare şi ea stă la baza întregii matematici moderne.

Definiţia 1.9.

Date două mulţimi A şi B spunem că am definit o relaţie binară între elementele mulţimii A şi elementele mulţimii B dacă dispunem de un procedeu (o regulă sau o modalitate) care să ne permită să decidem pentru orice pereche $(a, b) \in A \times B$ care din următoarele afirmaţii este adevărată:

a) perechea verifică relaţia,

b) perechea nu verifică relaţia.

Dacă relaţia o notăm cu R faptul că (a, b) verifică relaţia sau că a şi b se găsesc în relaţia R se scrie în forma aRb.

Mulţimea A se numeşte mulţime de pornire iar B mulţime de sosire.

Dacă R este o relaţie de la A la B atunci este evident că luînd toate perechile lui $A \times B$ o parte sau toate vor verifica relaţia. Să notăm tot cu R mulţimea respectivă adică:

$$R = \{(a, b) \in A \times B \mid aRb\}$$

Este clar că $R \subseteq A \times B$.

Prin urmare orice relaţie R de la A la B determină o submulţime a

produsului cartezian $A \times B$. Este valabilă şi afirmaţia reciprocă. Orice submulţime a produsului cartezian determină o relaţie de la A la B.

În adevăr fie $R \subseteq A \times B$. Să definim o relaţie R de la A la B în felul următor: $a R b$ dacă şi numai dacă $(a, b) \in R$. Această definiţie ne oferă regula (modalitatea) cu ajutorul căreia putem decide care dintre condiţiile a), b) este verificată pentru orice pereche (a, b). Aceste consideraţii ne permit să afirmăm că de la A la B se pot defini atîtea relaţii (există atîtea relaţii) cîte submulţimi are produsul cartezian $A \times B$ sau cîte elemente are mulţimea $\mathcal{P}(A \times B)$. Această strînsă legătură oferă posibilitatea să se definească şi altfel relaţia binară.

Definiţia 1.10.

Se numeşte relaţie binară definită de la A la B orice submulţime a produsului cartezian $A \times B$.

Într-o altă formulare se spune că o relaţie de la A la B este un triplet de forma (A, B, R) unde $R \subseteq A \times B$.

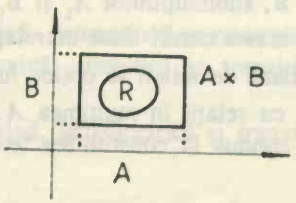
Mulţimea R se numeşte de obicei graful sau graficul relaţiei R. Conceptul de relaţie binară poate fi uşor generalizat prin conceptul de relaţie n-ară.

Definiţia 1.11.

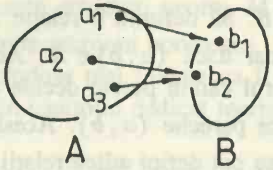
Se numeşte relaţie n-ară definită pe $A_1, \ldots, A_n$ orice submulţime a produsului cartezian $A_1 \times A_2 \times \ldots \times A_n$.

Pentru a obţine o formă mai intuitivă de prezentare a relaţiilor se folosesc mai multe modalităţi cum ar fi: carteziană, sagitală ş.a.

Prezentarea carteziană foloseşte reprezentarea carteziană în plan a produsului cartezian $A \times B$. De exemplu:



Prezentarea sagitală (cu săgeți) constă în a lega elementele $a \in A$ cu $b \in B$ care sînt în relație cu o săgeată de la a la b. De exemplu:

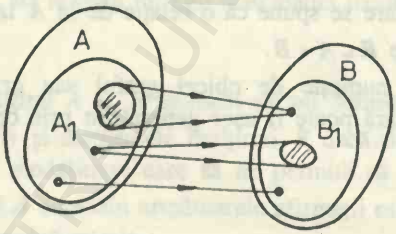


În acest caz $R = \{(a_1, b_1), (a_2, b_2), (a_3, b_2)\}$

Constatăm că o mulțime R definită de la A la B determină, în general, două relație $A_1 \subseteq A$ și $B_1 \subseteq B$ care se pot da prin:

$A_1 = \{a \in A \mid (\exists) b \in B \quad a,a \quad \hat{i}ncît \quad aRb\}$ și

$B_1 = \{b \in B \mid (\exists) a \in A \quad a,a \quad \hat{i}ncît \quad aRb\}$. Reprezentînd grafic obținem:



Analizînd reprezentarea precedentă constatăm că în general pentru relația R de la A la B în A pot exista elemente fără corespondent (cele din $A - A_1$), pot exista elemente cu un singur corespondent, pot exista mai multe elemente cu același corespondent și în sfîrșit pot exista elemente cu o mulțime de corespondenți (cu mai mult de un corespondent).

Individualizarea relațiilor se va putea face dacă vom impune condiții suplimentare mulțimilor A, B, submulțimilor A_1 și B_1 și regulilor de definire a relațiilor. Se vor obține în aceste cazuri clase de relații particulare.

O primă clasă importantă de relații se obține luînd $A = B$. În acest caz spunem că avem de a face cu relații în mulțimea A. Studierea problemelor puse de astfel de relații a condus la constituirea teoriei grafurilor care are

aplicaţii dintre cele mai importante în practică. Cu unele rezultate din această teorie vom lua cunoştinţă în capitolul dedicat acestei teorii.

Dacă relaţiile definite în mulţimea A îndeplinesc condiţii suplimentare se obţin relaţii speciale cum ar fi relaţia de echivalenţă, relaţia de ordine, relaţia de preordine etc., relaţii care au avut şi au un rol fundamental în edificarea matematicilor actuale.

Definiţia 1.12.

Vom spune că o relaţie R este relaţie funcţională dacă unui element din A îi corespunde cel mult un element din B. Cu alte cuvinte dacă un element din A are corespunzător atunci are numai unul *singur*.

Definiţia 1.13.

O relaţie funcţională o vom numi funcţie sau aplicaţie dacă $A_1 = A$, asta înseamnă că printr-o funcţie orice $a \in A$ are *un corespunzător unic în B*. Este de la sine înţeles că acest element din B poate fi acelaşi pentru mai multe elemente din A sau chiar pentru toate.

Se observă că în ultimul timp termenul cel mai des folosit este cel de aplicaţie, termenul de funcţie folosindu-se mai ales în cazurile în care A, B sînt mulţimi numerice. Să revenim la cazul general al relaţiilor de la A la A.

Dintre proprietăţile caracteristice ale relaţiilor definite într-o mulţime A amintim:

a) reflexivitatea $(\forall) a \in A, aRa$;

b) antireflexivitatea $(\forall) a \in A \quad a \not R a$

c) simetria $(\forall) a, b \in A, \quad aRb \to bRa$

d) antisimetria $(\forall) a, b \in A, \quad aRb$ şi $bRa \to a$ coincide cu b

e) tranzitivitatea $(\forall) a, b, c \in A, aRb$ şi $bRc \to aRc$

Definiţia 1.14.

1) O relaţie reflexivă şi tranzitivă se numeşte relaţie de *preordine*,

2) O relaţie reflexivă, simetrică şi tranzitivă se numeşte relaţie de *echivalenţă*,

3) O relaţie reflexivă, antisimetrică şi tranzitivă se numeşte relaţie de ordine.

Prezentăm în continuare cîteva consideraţii legate de relaţiile de ordine şi echivalenţă.

Ideea de ordine este dintre cele mai importante din matematică. Ea constituie aprofundarea noţiunii intuitive de ordine (ierarhizare) prin care se compară elementele unei perechi în scopul stabilirii care este primul şi care este următorul în funcţie de un anumit criteriu.

Aşa cum am mai spus o relaţie de ordine este reflexivă, antisimetrică şi tranzitivă. De regulă relaţia de ordine sau ordonarea este indicată prin simbolurile : $\leq$, $\subseteq$, $\preceq$, ş.a.

O mulţime în care există o relaţie de ordine se spune că este înzestrată cu o *structură de ordine* sau pur şi simplu că este o mulţime ordonată. Pentru a evidenţia că o mulţime A este ordonată prin relaţia „$\leq$" se scrie $(A, \leq)$.

O mulţime cu structura de ordine $(A, \leq)$ se numeşte *total ordonată* dacă $(\forall) a, b \in A$ urmează $a \leq b$ sau $b \leq a$. Dacă nu se face precizarea că fiind date două elemente oarecare $a, b \in A$ avem $a \leq b$ sau $b \leq a$ se spune că avem o ordine parţială iar mulţimea A este *parţial ordonată*.

Într-o structură de ordine dacă $a \leq b$ sau $b \leq a$ se spune că a, b sînt *comparabile*.

Dacă $(A, \leq)$ este parţial ordonată şi $S \subseteq A$ atunci evident că în S putem pune problema ordonării induse de ordinea din A. Adică $x \leq y$ în S dacă $x \leq y$ în A.

Evident că această relaţie indusă este tot o parţială ordonare şi deci orice submulţime a unei mulţimi parţial ordonate este parţial ordonată.

Un element $a \in (A, \leq)$ se numeşte prim element (ultim element) al lui A dacă $a \leq x$ $(x \leq a)$, $\forall x \in A$. Aceste elemente sînt unice şi se notează cu 0 respectiv 1. Luînd $S \subseteq A$ se impun clarificate încă unele elemente speciale.

Elementul $a \in A$ se numeşte *superiorul* (*inferiorul*) mulţimii S dacă $(\forall) x \in S$, $x \leq a$ şi dacă există $a' \in A$ aşa ca $x \leq a'$, atunci $a \leq a'$ (respectiv $a \leq x$ şi dacă $a' \leq x$ atunci $a' \leq a$).

Elementul a (dacă există) se notează cu sup. (S) (respectiv inf. (S)).

O mulţime parţial ordonată $(A, \leq)$ se numeşte *latice* dacă pentru orice două elemente $a, b \in A$ există superiorul şi inferiorul. Se notează $inf\ \{a, b\} = a \cap b$, $sup\ \{a, b\} = a \cup b$ iar structura corespunzătoare $(A, \cap, \cup)$.

În fine dacă orice submulţime S din $(A, \leq)$ are un prim element se zice că A este *bine ordonată*.

De exemplu $(\mathbb{N}, \leq)$ este total ordonată în raport cu ordinea naturală, $(\mathcal{P}(M), \leq)$ parţial ordonată în raport cu incluziunea. Pentru relaţia de echivalenţă se impun cîteva consideraţii aparte.

Fie $\mathbb{R}$ o relaţie de echivalenţă în mulţimea A. Două elemente $a, b \in A$ în relaţia $\mathbb{R}$ se numesc asociate sau ehivalente.

Se numeşte clasă de echivalenţă a elementului $a \in A$ şi o vom nota cu C_a sau $\hat{a}$ mulţimea tuturor elementelor $x \in A$ echivalente cu a, deci:

$$C_a = \hat{a} = \{b \mid b \in A \ , \ b \mathbb{R} a\}$$

În loc de $a \mathbb{R} b$ se mai obişnuieşte să se scrie $a \sim b \pmod{\mathbb{R}}$ prin analogie cu o notaţie din teoria numerelor.

Teorema 1.1.

Fiind dată o relaţie de echivalenţă $\mathbb{R}$ în mulţimea A, clasele de echivalenţă constituie o partiţie a lui A şi invers, dată o partiţie a lui A aceasta determină o relaţie de echivalenţă în A.

Demonstraţie: Pentru a demonstra prima parte a teoremei să arătăm că două clase sînt sau disjuncte sau coincid, că orice clasă este nevidă şi că reuniunea claselor este tocmai A.

Este evident că $C_a \neq \Phi$ deoarece $a \in C_a$, $\mathbb{R}$ fiind reflexivă. Fie $m \in A$ şi $m \in C_a \cap C_b$. Urmează că $m \in C_a$ şi $m \in C_b$ şi deci $m \mathbb{R} a$ şi $m \mathbb{R} b$. Să arătăm că $C_a = C_b$. Fie $x \in C_a$ atunci $x \mathbb{R} a$ şi cum $a \mathbb{R} m$ rezultă $x \mathbb{R} m$ şi apoi avînd în vedere că $m \mathbb{R} b$, putem spune că $x \mathbb{R} m$ adică $x \in C_b$. Aşa dar $C_a \subseteq C_b$. Analog se arată că $C_b \subseteq C_a$. Deci dacă două clase au măcar un element comun atunci ele coincid. Se vede deasemeni că:

$$\bigcup_{a \in A} C_a = A$$

Să arătăm acum şi a doua parte a teoremei. Fie $A_1, A_2, \ldots, A_n$ o partiţie a mulţimii A; înseamnă că:

1. $(\forall) i \quad A_i \neq \Phi$

2. dacă $i \neq j \quad A_i \cap A_j = \Phi$

3. $A = \bigcup_{i=1}^{n} A_i$

Să definim în A o relaţie $\mathbb{R}$ astfel: $a \mathbb{R} b$ atunci şi numai atunci cînd $(\exists) i \quad i = \overline{1,n}$ aşa încît $a, b \in A_i$. Relaţia $\mathbb{R}$ este aşa cum se poate vedea imediat, o relaţie de echivalenţă fiind reflexivă simetrică şi tranzitivă. Clasele acestei relaţii sînt tocmai submulţimile din familia dată care forma iniţial partiţia. Mulţimea care are drept elemente clasele de echivalenţă determinate de $\mathbb{R}$ se numeşte mulţimea cît sau mulţime factor a lui A prin $\mathbb{R}$ şi se notează cu $A/\mathbb{R}$.

Din cele precedente urmează că într-o mulţime putem defini atîtea relaţii de echivalenţă cîte partiţii admite A. Relaţiile de echivalenţă sînt de importanţă capitală pentru întreaga algebră modernă.

Aşa cum am văzut deja o clasă importantă de relaţii sînt relaţiile funcţionale cu proprietatea că oricare ar fi $a \in A$ există un b unic în B aşa încît $a \mathbb{R} b$. Aceste relaţii le-am numit *funcţii* sau aplicaţii. Deci o funcţie asociază *fiecărui* $a \in A$ un b şi numai unul din B. Aplicaţiile le vom nota cu f, g,...

A defini o aplicaţie f a mulţimii A în mulţimea B revine deci la a preciza pentru fiecare $x \in A$ corespunzătorul din B pe care-l notăm cu $f(x)$. Se scrie $f: A \to B$.

Un element oarecare al mulţimii A se mai numeşte variabilă sau argument al aplicaţiei. Dacă $f: A \to B$ şi $a \in A$, elementul $f(a)$ se numeşte imaginea sa sau valoarea lui f în a.

Fie două aplicaţii $f: A \to B$ şi $g: C \to D$. Este natural să spunem că cele două aplicaţii sînt egale dacă $A = C$, $B = D$ şi $(\forall) x \in A$, $f(x) = g(x)$.

Pornind de la relaţii am ajuns la noţiunea de funcţie studiată în liceu şi din acest motiv nu vom mai face o detaliere a chestiunilor legate de proprietăţi, clasificări etc.

Observaţia 1. Orice aplicaţie $f: A \to B$ determină o relaţie de echivalenţă pe A.

În adevăr să definim cu ajutorul lui f o relaţie: R_f de la A la A în felul

următor: $\forall a, b \in A$ să punem prin definiţie $a R_f b \leftrightarrow f(a) = f(b)$. Deci două elemente din A sînt în relaţia R_f dacă şi numai dacă au aceeaşi imagine prin f. Relaţia R_f este o relaţie de echivalenţă. În adevăr avem: $f(a) = f(a)$; dacă $f(a) = f(b)$ atunci $f(b) = f(a)$ şi în sfîrşit dacă $f(a) = f(b)$ şi $f(b) = f(c)$ atunci $f(a) = f(c)$. Prin urmare R_f este reflexivă, simetrică şi tranzitivă adică este echivalentă pe A.

Observaţia 2. Legăturile dintre o aplicaţie $f: A \to B$ şi rezultatele operaţiilor între submulţimi din A şi B sînt ilustrate de următoarele relaţii (formule) pe care le propunem spre demonstrarea cititorului:

Dacă X, Y, submulţimi din A au loc:

1. $X \subseteq Y$, urmează $f(X) \subseteq f(Y)$;
2. $f(X \cup Y) = f(X) \cup f(Y)$;
3. $f(X \cap Y) \subseteq f(X) \cap f(Y)$ şi dacă f este injectivă $f(X \cap Y) = f(X) \cap f(Y)$

Observaţia 3. Fie A, B, C, mulţimi nevide date şi să considerăm $M \subseteq A \times B$ şi aplicaţia $f: M \to C$.

Definiţia 1.15.

Orice aplicaţie f, din cele de mai sus, se numeşte *operaţie* definită în $A \times B$ cu valori în C.

Dacă $M = A \times B$ spunem că operaţia este peste tot definită în $A \times B$ în caz contrar spunem că este parţial definită.

Operaţia o vom numi *operaţie algebrică* sau lege de compoziţie în cazurile $C \subseteq A$ sau $C \subseteq B$.

Aşa dar putem vorbi de operaţii algebrice peste tot definite sau parţial definite pe $A X B$.

Dacă $A = B = C$ spunem că avem de a face cu o lege de compoziţie binară internă.

Dacă $A = \underbrace{B \times \ldots \times B}_{(n-1)-ori}$ şi $B = C$ operaţia sau legea de compoziţie definită în B o vom numi operaţie n-ară internă.

În fine dacă $A \neq B$ dar $B = C$ legea de compoziţie o vom numi lege de compoziţie externă pe B cu operatori din A. Sînt multe situaţii în care unele operaţii interne sînt interpretate ca operaţii externe; în asemenea situaţii atît elementele cît şi operatorii aparţin aceleiaşi mulţimi.

Prin urmare o operaţie internă peste tot definită în A este o aplicaţie $f: \underbrace{A \times \ldots \times A}_{n-ori} \to A$, iar o operaţie externă pe A o aplicaţie $f: K \times A \to A$ sau $f: A \times K \to A$.

În general se lucrează cu operaţii peste tot definite însă nici cele parţial definite nu trebuiesc neglijate deoarece au jucat şi ele rol important în teoria generală.

Definiţia 1.16.

O mulţime A în care avem definite o mulţime Ω de operaţii algebrice interne sau externe se numeşte *structură algebrică, algebră universală* sau *Ω-algebră*. Deoarece o mulţime poate fi structurată algebric în foarte multe moduri s-a impus necesitatea ca atunci cînd notăm structura să punem mulţimea după care să trecem semnul operaţiilor avute în vedere; de exemplu $(N, +)$, $(N, \cdot)$, $(Z, +, \cdot)$ etc.

Se cuvine să atragem atenţia că studiul structurilor algebrice, în cazul cel mai general pe care l-am considerat, nu poate fi făcut. Din acest motiv se are în vedere faptul că legile de compoziţie care determină o anume structură de obicei au anumite proprietăţi. Aceste proprietăţi individualizează structura sau clasa din care face parte o anumită structură. Structurile algebrice studiate în liceu sînt edificatoare în acest sens.

Studiul structurilor algebrice ar fi fost sărac dacă nu s-ar fi considerat şi aplicaţii între structuri. Dintre toate aceste aplicaţii interesează cele care conservă structurile, sînt compatibile cu operaţiile şi care au fost numite *morfisme*.

Definiţia 1.17.

Dacă $(A, *)$ şi (B, ν) sînt două structuri algebrice atunci un morfism între aceste structuri este o aplicaţie $f: A \to B$ cu proprietatea:

$$f(a * b) = f(a) \nu f(b) \quad (\forall) a, b \in A$$

§ 4. ELEMENTE DE ANALIZĂ COMBINATORIE

Materia de liceu pune la îndemîna elevilor primele cunoştinţe de combinatorică. Acestea sînt însă insuficiente şi maniera de introducere frumoasă dar abstractă nu are darul să formeze o imagine unitară asupra problemelor de acest tip.

Folosirea unor variante diferite pentru definirea noţiunilor fundamentale a condus la convingerea că aranjamentele sînt aplicaţii iar combinările submulţimi.

În fond, istoric vorbind, analiza combinatorie s-a constituit prin studierea problemelor legate de cîte combinaţii, aranjamente, în unele sau alte condiţii se pot forma cu elemente dintr-o mulţime dată. În majoritatea lor problemele de început au fost probleme în care se cerea numărarea diferitelor modalităţi de grupare şi ordonare a unor elemente după reguli date. Alegere şi şansă iată preocupările iniţiale care au impulsionat căutările în acest domeniu.

Dacă la început toate cercetările erau orientate spre probleme stranii apărute în special în jocurile de noroc, încetul cu încetul preocupările se mută spre profunde probleme de matematică finită. Numeroase probleme din: teoria probabilităţilor, fundamentele geometriei, statistică matematică şi multe alte domenii s-au grupat dînd un capitol mare al matematicii moderne numit combinatorică.

Metodele combinatoricii sînt deosebit de diversificate şi rezultatele de mare profunzime şi aplicabilitate. În prezent metodele combinatoricii se folosesc pentru rezolvarea problemelor de transport, pentru stabilirea planurilor de producţie, în statistică etc. Din aceste motive am considerat potrivit să revenim supra elementelor iniţiale pentru a valorifica mai bine modul clasic de punere a problemelor şi de găsire a răspunsurilor.

Pentru stabilirea formulelor sînt suficiente următoarele două reguli:

a) *Regula sumei*. Dacă mulţimea A are m elemente şi mulţimea B are n elemente şi dacă cele două mulţimi sînt disjuncte atunci $A \cup B$ are $m + n$ elemente.

În multe cazuri această regulă se formulează şi sub forma „Dacă obiectul A poate fi ales în m moduri şi obiectul B în alte n moduri atunci „sau A sau B" poate fi ales în $m + n$ moduri".

b) *Regula produsului*. Dacă mulţimea A are m elemente iar mulţimea

B are n elemente atunci $A \times B$ are $m \cdot n$ elemente. În altă formă această regulă se mai prezintă şi astfel: „Dacă obiectul A poate fi ales în m moduri şi după fiecare alegere, obiectul B poate fi ales la rîndul său în n moduri atunci „A şi B" se poate alege în $m \cdot n$ moduri".

Definiţia 1.18.

a) Se numeşte o *aranjare* de m obiecte din n obiecte date o alegere ordonată a r din cele n obiecte.

b) Se numeşte o *combinare* de r obiecte din n obiecte date o alegere a r din cele n obiecte fără a ţine seama de ordine.

În legătură cu aceste definiţii trebuie să facem următoarele observaţii: mai întîi nici una din ele nu se vorbeşte nimic despre caracterul celor n elemente: ele pot fi toate de acelaşi fel, unele de un fel altele de altul sau toate diferite. Cel mai general caz este următorul: există k feluri de elemente iar de felul j, n_j elemente aşa încît:

$$n_1 + n_2 + \ldots + n_k = n$$

În al doilea rînd cuvîntul „ordonat" înseamnă că două alegeri se consideră diferite chiar dacă au aceleaşi elemente însă acestea sînt aşezate în ordini diferite. Folosind regulile de mai sus se pot găsi cu uşurinţă formulele de calcul pentru toate cazurile:

a) *Elementele sînt diferite.*

Situaţia se prezintă astfel: avem n obiecte distincte şi vrem să vedem în cîte moduri putem alege grupe ordonate de cîte n obiecte. Se poate raţiona astfel: primul din cele r obiecte poate fi ales în n moduri, al doilea poate fi ales în $(n-1)$ moduri,..., obiectul de pe locul r în $n-(r-1)$ moduri.

Dacă notăm numărul aranjărilor cu A_n^r şi aplicăm regula produsului se obţine:

$$A_n^r = n(n-1)\ldots(n-r+1)$$

În cazul în care $r = n$ aranjările le vom numi *permutări* iar numărul lor îl notăm cu $\mathcal{P}_n$. Deci:

$$\mathcal{P}_n = n(n-1)\ldots 3 \cdot 2 \cdot 1 = n!$$

Deoarece factorialele sînt tabelate este utilă şi formula:

$$A_n^r = \frac{n!}{(n-r)!}$$

b) *Elementele nu sînt diferite.*

Cele n obiecte sînt de mai multe feluri. Astfel p sînt de un fel, q de altul, r de altul ş.a.m.d. În acest caz se consideră grupe care conţin toate cele n obiecte (elemente). Datorită faptului că unele elemente se repetă grupele au căpătat denumirea de *permutări cu repetiţie.*

Fie x numărul acestor grupe. Pentru a putea găsi valoarea lui x să presupunem că cele p elemente de primul fel le-am înlocuit cu elemente distincte. Acestea le putem schimba între ele în $p!$ moduri. Prin urmare numărul acestor noi grupe este $x \cdot p!$. Procedînd la fel cu toate grupele vom înlocui toate elementele cu elemente distincte iar numărul lor va fi pe de o parte: $xp!q!r!....$ Pe de altă parte aceste grupe sînt tocmai permutările din n obiecte, deci numărul lor este $n!$. Am obţinut deci: $xp!q!r!... = n!$ şi deci:

$$x = \frac{n!}{p!q!r!...}$$

c) *Cazul în care avem n feluri de obiecte iar în fiecare fel avem un număr nelimitat de elemente.*

Dorim să facem grupe de cîte r obiecte din cele n feluri, feluri fel putîndu-se repeta pînă la r ori. Astfel de grupuri le vom numi *aranjări cu repetiţie*, iar numărul lor se notează cu a_n^r. Raţionînd ca şi la aranjamentele simple obişnuite vom observa că: primul obiect poate fi ales în n moduri al doilea în n moduri,..., obiectul de pe locul r în n moduri (deoarece din fiecare fel avem un număr nelimitat de elemente). Aplicînd regula produsului se obţine imediat:

$$a_n^r = n^r$$

d) *Combinări de n obiecte luate cîte r*

În acest caz ordinea nu este luată în considerare. Să notăm cu C_n^r sau $\binom{n}{r}$ numărul acestor grupe. Considerîndu-le ca grupe ordonate, din fiecare grupă se obţin $r!$ grupe noi. Cum noile grupe sînt tocmai aranjări rezultă că: $r!\, C_n^r = A_n^r$ de unde

$$C_n^r = \frac{A_n^r}{r!}$$

Avînd în vedere expresia lui A_n^r obţinem:

$$C_n^r = \frac{n!}{r!\,(n-r)!}$$

e) *Combinări cu repetiţie.*

Avem drept scop să numărăm numărul grupelor, combinărilor din n obiecte luate cîte r, fiecare din ele putîndu-se repeta de 0, 1, 2, ... pînă la r ori. Să notăm cu K_n^r. Pentru a găsi formula de calcul vom folosi o metodă indicată de Euler. Fie $c_1, c_2, \ldots, c_r$ o astfel de grupă. Cum natura elementelor nu are importanţă rezultă că putem să le înlocuim cu numere, şi cum nici ordinea nu contează să le punem în ordine crescătoare:

$$c_1 \le c_2 \le c_3 \le \ldots \le c_r$$

Asociem fiecărui asemenea şir cu elemente ce se pot repeta un nou şir obţinut astfel: $d_1 = c_1 + 0$, $d_2 = c_2 + 1, \ldots, d_r = c_r + r - 1$. Grupa $d_1 d_2 \ldots d_r$ va avea toate elementele distincte. Prin urmare fiecărei grupe iniţiale cu elemente care se repetă îi putem araşa o grupă unică bine determinată. Evident că această corespondenţă este injectivă şi deci vom putea să numărăm cîte grupe am avut iniţial numărîndu-le pe cele obţinute. Grupele obţinute, însă, sînt combinări obişnuite de $n + r - 1$ obiecte luate cîte r. Aşa dar:

$$K_n^r = C_{n+r-1}^r$$

Procedînd în acest mod putem regăsi toate formulele de care avem nevoie în procesele de numărare.

Întrebări de autocontrol, exerciţii

1. Puteţi defini noţiunea de mulţime?
2. Prezentaţi prin extensie şi comprehensiune mulţimea $\mathcal{P}$ a numerelor pare.
3. Considerînd $M = \{0,1,2,3,4,5,7,8,9\}$ se cere să se formeze mulţimea $\mathcal{P}(M)$. Cîte elemente are $\mathcal{P}(M)$?
4. Fie $A = \{1,2,3,4,5\}$. Reprezentaţi cartezian şi sagital relaţiile binare definite prin $x < y$, $|x - y| = 2$ şi $|x - 2| = |2y - 3|$
5. Fie $A = \{a,b\}$ şi $B = \{1,2,3\}$. Cîte funcţii $f: A \to B$ există?
6. Daţi exemple de relaţii de ordine şi de relaţii de echivalenţă.
7. Rezolvaţi ecuaţia $K_7^x = 28$.

BCU IAŞI/CENTRAL UNIVERSITY LIBRARY

CAPITOLUL II

ELEMENTE DE ALGEBRĂ

§ 1. SPAŢII VECTORIALE (SPAŢII LINIARE)

Fie $S \neq \varnothing$ şi k un corp comutativ.

Definiţia 2.1.

Spunem că S este un k-spaţiu vectorial dacă pe S este definită o operaţie internă binară notată $+$ faţă de care este grup abelian şi există o operaţie externă, prin care $(\forall)\ \alpha \in k$ şi $u \in S$ îi asociem elementul αu din S şi care îndeplineşte condiţiile:

1. $\alpha(u+v) = \alpha u + \alpha v$
2. $(\alpha + \beta)u = \alpha u + \beta u$. + grup abelian
3. $\alpha(\beta u) = (\alpha\beta)u$
4. $1 \cdot u = u$

pentru $1, \alpha, \beta, \ldots$ din k şi $u, v, \ldots$ din S.

Elementele lui S se numesc vectori, iar operaţia din S adunarea vectorilor.

Operaţia externă se numeşte produs cu scalari. Elementul nul din S şi elementul nul din k le notăm cu 0.

Consecinţe:

1. Dacă S este k-spaţiu atunci $(S, +)$ este grup abelian.

BCU IASI / CENTRAL UNIVERSITY LIBRARY

Într-adevăr fie $u,v \in S$: Putem scrie:

$$(u+v)+(u+v) = 2(u+v) = 2u+2v = u+u+v+v$$

Ţinînd cont că S este grup se obţine:

$$v+u = u+v$$

pentru orice $u,v \in S$

2. Dacă $\alpha u = 0$ atunci $\alpha = 0$ sau $u = 0$.

Mai întîi observăm că $0u = (0+0)u = 0u+0u$ şi deci $0u = 0$. Apoi $\alpha 0 = \alpha(0+0) = \alpha 0 + \alpha 0$ şi deci şi $\alpha 0 = 0$.

Pentru partea a doua avem: $\alpha u = 0$.. Dacă $\alpha \neq 0$ rezultă că există $\alpha^{-1} \in k$ şi deci $\alpha^{-1}(\alpha u) = 1 \cdot u = u$. Adică $u = 0$.

3. $(\forall)$ $\alpha \in k$ şi $u \in S$ avem:

$$(-\alpha)u = \alpha(-u) = -(\alpha u)$$

$$(-\alpha)(-u) = \alpha u$$

Într-adevăr $0u = 0 \Rightarrow [\alpha + (-\alpha)]u = \alpha u + [(-\alpha)u] = 0$. Dar $\alpha u + (-(\alpha u)) = 0$ şi deci $(-\alpha)u = -(\alpha u)$. Apoi $\alpha \cdot 0 = 0 \Rightarrow$ $\Rightarrow \alpha(u+(-u)) = 0 \Rightarrow \alpha u + \alpha(-u) = 0$ şi mai avem $\alpha u + (-(\alpha u)) = 0$ şi deci $\alpha(-u) = -\alpha u$. Pentru partea a doua avem:

$$(-\alpha)(-u) = -(\alpha(-u)) = -(-\alpha u) = \alpha u$$

4. $(\forall)$ $\alpha, \beta \in k$ şi $u,v \in S$ avem:

$$(\alpha - \beta)u = \alpha u - \beta u$$

$$\alpha(u-v) = \alpha u - \alpha v$$

Avînd în vedere precedentele putem scrie succesiv:

$$(\alpha - \beta)u = \alpha u + (-\beta)u = \alpha u + (-(\beta u)) = \alpha u - \beta u$$

$$\alpha(u-v) = \alpha u + \alpha(-v) = \alpha u + (-(\alpha v)) = \alpha u - \alpha v$$

BCU IASI / CENTRAL UNIVERSITY LIBRARY

5. Dacă $\alpha, \alpha_1, \alpha_2, \ldots, \alpha_n \in k$, $u, v_1, v_2, \ldots v_m \in S$ şi $n, m \in \mathbb{N}$ au loc relaţiile:

$$(\alpha_1 + \ldots + \alpha_n)u = \alpha_1 u + \ldots + \alpha_n u$$

$$\alpha(u_1 + \ldots + u_m) = \alpha u_1 + \ldots + \alpha u_m$$

$$(\alpha_1 + \ldots + \alpha_n)(u_1 + \ldots + u_m) = \alpha_1 u_1 + \ldots + \alpha_n u_m$$

Aceste relaţii se scriu prescurtat:

$$(\sum_{i=1}^{n} \alpha_i)u = \sum_{i=1}^{n} (\alpha_i u), \alpha(\sum_{j=1}^{m} u_j) = \sum_{j=1}^{m} (\alpha u_j)$$

$$(\sum_{i=1}^{n} \alpha_i)(\sum_{j=1}^{m} u_j) = \sum_{i=1}^{n} \sum_{j=1}^{m} \alpha_i u_j$$

Toate aceste trei relaţii se demonstrează cu uşurinţă folosindu-se metoda inducţiei complete.

Exemple:

1. Mulţimea $M_{n\times m}(\mathbb{R})$ este un $\mathbb{R}$-spaţiu vectorial în raport cu operaţiile obişnuite.

2. Mulţimea $k[X]$ a polinoamelor cu coeficienţii din k este un k-spaţiu. Pentru $k = \mathbb{R}, \mathbb{C}$, se obţin spaţii deosebit de interesante.

3. Mulţimea $k^n = k \times k \times \ldots \times k$ este un k-spaţiu dacă $(\alpha_1, \ldots, \alpha_n) + (\beta_1 + \ldots + \beta_n) = (\alpha_1 + \beta_1, \ldots, \alpha_n + \beta_n)$ şi $\alpha(\alpha_1, \ldots, \alpha_n) = (\alpha\alpha_1, \ldots, \alpha\alpha_n)$.

În particular pentru $k = \mathbb{R}$ sau $k = \mathbb{C}$ se obţin spaţiile $R^n, \mathbb{C}^n$. Trebuie să observăm că avem $k^n = M_{1\times n}(k)$ şi deci putem spune că elementele acestor spaţii sînt matrici cu o linie şi n coloane.

4. Spaţiul nul. Dacă notăm cu O mulţimea cu un singur element notat cu 0 atunci definind $0 + 0 = 0$ şi $(\forall)\ \alpha \in k\ \ \alpha \cdot 0 = 0$ putem spune că O este k-spaţiu. Acest spaţiu se numeşte spaţiu nul.

Atenţie, spaţiul nul are elemente.

5. Să notăm cu $k_n[X]$ mulțimea polinoamelor cu coeficienți din corpul k și de grad mai mic sau egal cu n. În raport cu operațiile obișnuite $k_n[X]$ este un k-spațiu vectorial.

6. Orice corp k poate fi interpretat ca un k-spațiu vectorial în raport cu adunarea și înmulțirea lui k.

§ 2. SUBSPAȚII LINIARE

Definiția 2.2.

O submulțime nevidă din k-spațiul S se numește k-subspațiu vectorial (sau simplu subspațiu) dacă este k-spațiu în raport cu operațiile induse.

Teorema 2.1.

Fie $S_1 \neq \varnothing$, $S_1 \subseteq S$. Următoarele afirmații sînt echivalente:

(1) S_1 subspațiu în S;

(2) $(\forall)\ u, v \in S_1,\ (\forall)\ \alpha \in k \to u - v \in S_1, \alpha u \in S_1$;

(3) $(\forall)\ u, v \in S_1,\ (\forall)\ u + v \in S_1, \alpha u \in S_1$;

(4) $(\forall)\ u, v \in S_1,\ (\forall)\ \alpha, \beta \in k \to \alpha u + \beta v \in S_1$;

Demonstrație: Să arătăm că (1) $\leftrightarrow$ (2).

Dacă S_1 subspațiu în $S \to S_1$ subgrup în grupul aditiv S și deci $(\forall)\ u, v \in S_1 \to u - v \in S_1$ și de asemenea evident și $\alpha u \in S_1$ pentru orice $\alpha \in k$.

Reciproc: O submulțime $S_1 \neq \varnothing$ cu proprietatea (2) este evident subgrup în grupul aditiv $(S, +)$ stabilă la operația externă. Restul axiomelor din definiție se verifică cu ușurință. Analog se justifică și cu celelalte echivalențe.

Observație: O și S sînt subspații în k-spațiul S și se numesc subspații improprii sau subspații banale. Un interes deosebit prezintă subspațiile nebanale.

Exemple:

1. $k_n[X]$ este subspațiu în $k[X]$.

2. Fie $u_1, u_2, \ldots, u_n$ vectori din k-spațiul S.

Definiţia 2.3.

Vectorul $u \in S$ este o combinaţie liniară de vectori $u_1, u_2, \ldots, u_n$ dacă există $\alpha_1, \alpha_2, \ldots, \alpha_n \in k$ astfel încît u se poate scrie în forma:

$$u = \alpha_1 u_1 + \alpha_2 u_2 + \ldots + \alpha_n u_n$$

Să notăm mulţimea combinaţiilor liniare de $u_1, u_2, \ldots, u_n$ cu simbolul $<u_1, u_2, \ldots, u_n>$.

Mulţimea $<u_1, u_2, \ldots, u_n>$ este un subspaţiu liniar în S numit subspaţiul generat de $u_1, \ldots, u_n$ iar fiecare u_i se numeşte generator.

Justificarea afirmaţiei rezultă imediat din faptul că are loc T 2.1 cazul (3). De aici se obţine o sursă foarte importantă de subspaţii ale unui spaţiu dat (dacă acestea există). Se spune că spaţiul S este finit generat dacă există o submulţime finită de vectori aşa fel încît subspaţiul generat de ei să fie chiar spaţiul S.

În toate consideraţiile care urmează vom admite că spaţiile respective sînt finit generate. Dacă S_1, S_2 sînt subspaţii ale spaţiului S atunci în multe cazuri se pune problema operaţiilor cu ele.

Definiţia 2.4.

a) Se numeşte intersecţia subspaţiilor S_1 şi S_2 mulţimea:

$$S_3 = S_1 \cap S_2$$

b) Se numeşte suma subspaţiilor S_1 şi S_2 mulţimea :

$$S_4 = S_1 + S_2 = \{x \in S / x = x_1 + x_2, x_1 \in S_1, x_2 \in S_2\}$$

Teorema 2.2.

Intersecţia şi suma a două subspaţii este tot un subspaţiu în S.

Demonstraţie: Observăm că S_3 şi S_4 sînt mulţimi nevide deoarece măcar vectorul nul, care face parte din toate subspaţiile, face parte şi din S_3 şi din S_4.

Fie $u, v \in S_3$ şi $\alpha, \beta \in k \Rightarrow \alpha u \in S_1$, $\alpha u \in S_2$ şi analog $\beta v \in S_1, S_2$ şi deci deoarece S_1, S_2 subspaţii conform T2.1 şi $\alpha u + \beta v \in S_1, S_2$ adică $\alpha u + \beta v \in S_3$.

Pentru S_4 procedăm astfel. Luăm $u, v \in S_4$ atunci există $u_1, v_1 \in S_1$ şi $u_2, v_2 \in S_2$ astfel încît $u = u_1 + v_1$ şi $v = u_2 + v_2$. În acest caz pentru orice $\alpha, \beta \in k$ avem:

$$\alpha u + \beta v = \alpha(u_1 + v_1) + \beta(u_2 + v_2) = (\alpha u_1 + \beta v_1)(\alpha u_2 + \beta v_2) = u_3 + v_3$$

unde $u_3 \in S_1$, $v_3 \in S_2$.

Deci $\alpha u + \beta v \in S_4$ şi conform T2.1 punctul (4) S_4 este subspaţiu.

Observaţie: S_3 este chiar intersecţia de mulţimi dar S_4 nu coincide cu reuniunea mulţimilor S_1 şi S_2. Subspaţiul sumă este „ cel mai mic" subspaţiu care conţine subspaţiile S_1 şi S_2.

Dacă $S_1 \cap S_2 = 0$ atunci subspaţiul sumă S_4 se numeşte suma directă a subspaţiilor S_1 şi S_2 şi se notează:

$$S_4 = S_1 \oplus S_2$$

Dacă $S_1 \oplus S_2 = S$ atunci S_2 se numeşte subspaţiul complementar lui S_1 în S.

§ 3. DEPENDENŢĂ LINIARĂ, BAZE, COORDONATE

Să revenim acum din nou la familii de vectori.

Definiţia 2.5.

Fie $u_1, u_2, \ldots, u_n \in S$

a) O relaţie de forma (1) $\alpha_1 u_1 + \alpha_2 u_2 + \ldots + \alpha_n u_n = 0$, $\alpha_i \in k$ se numeşte relaţie de dependenţă liniară.

b) Mulţimea $u_1, \ldots, u_n$ este liniar dependentă dacă (1) are loc cu măcar un $\alpha_i \neq 0$.

c) Vectorii $u_1, \ldots, u_n$ se numesc liniar independenţi dacă relaţia (1) are loc numai dacă toţi $\alpha_i = 0$.

Consecinţe:

1) Vectorul nul este totdeauna liniar dependent deoarece $1 \cdot 0 = 0$

2) Orice vector $u \neq 0$ este liniar independent deoarece

$\alpha u = 0$ şi $u \neq 0 \Rightarrow \alpha = 0$

3) O familie (o mulţime) de vectori din S care conţine o submulţime liniar dependentă este liniar dependentă. În adevăr fie $u_1,...,u_n$ cu $u_1,...,u_m$ $(m<n)$ liniar dependentă. Atunci $\alpha_1 u_1 + ... + \alpha_m u_m = 0$ are loc cu măcar un $\alpha_i \neq 0$. Această relaţie o putem scrie $\alpha_1 u_1 + ... + \alpha_m u_m + 0u_{m+1} + ... + 0u_n = 0$ relaţie care arată liniara dependenţă a vectorilor $u_1,...u_n$. Urmează, de asemenea, că orice familie de vectori care conţine vectorul nul este liniar dependent.

4) Vectorii $u_1,...u_n$, $n>1$ sînt liniar dependenţi dacă şi numai dacă unul dintre ei se exprimă ca o combinaţie liniară de ceilalţi. În adevăr dacă $u_1,...,u_n$ sînt liniar independenţi atunci $\alpha_1 u_1 + \alpha_2 u_2 + ... + \alpha_n u_n = 0$ are loc cu măcar un $\alpha_i \neq 0$. Fie de exemplu $\alpha_n \neq 0$. Atunci $u_n = -\frac{\alpha_1}{\alpha_n} u_1 - \frac{\alpha_2}{\alpha_n} u_2 - ... - \frac{\alpha_{n-1}}{\alpha_n} u_{n-1}$

adică $u_n \in <u_1, u_2, ..., u_{n-1}>$.

Reciproc. Dacă $u_i = \alpha_1 u_1 + ... + \alpha_{i-1} u_{i-1} + \alpha_{i+1} u_{i+1} + ... + \alpha_n u_n$ avem $\alpha_1 u_1 + ... + \alpha_{i-1} u_{i-1} + (-1) u_i + \alpha_{i+1} u_{i+1} + ... + \alpha_n u_n = 0$. De aici urmează evident că $u_1,...,u_n$ liniar dependenţi.

5) Dacă mulţimea de vectori $\{u_1,...,u_n\}$ este liniar independentă atunci orice submulţime a ei conţine tot vectori independenţi. În adevăr fie vectorii $u_1,...,u_m$ $(m<n)$ liniar dependenţi. Atunci există relaţia $\alpha_1 u_1 + ... + \alpha_m u_m = 0$ cu măcar un $\alpha_i \neq 0$. Această relaţie se mai poate scrie:

$$\alpha_1 u_1 + \alpha_2 u_2 + ... + \alpha_m u_m + 0u_{m+1} + ... + 0u_n = 0$$

De aici rezultă $u_1,...,u_n$ liniar dependenţi în contradicţie cu ipoteza.

Definiţia 2.6.

Un sistem (o mulţime) de vectori $u_1,...,u_n$ din spaţiul S se numeşte bază dacă:

1) $u_1,...,u_n$ sînt liniar independenţi

2) $<u_1,...,u_n> = S$

Existenţa unei baze într-un spaţiu finit generat o presupunem fără demonstraţie.

Fie $u_1,...,u_n$ o bază în k-spaţiul S şi $u \in S$ atunci $u,u_1,u_2,...,u_n$ sînt liniar dependenţi şi deci avem

(2) $u = \alpha_1 u_1 + ... + \alpha_n u_n$

Vom arăta acum că scalarii $\alpha_1,...,\alpha_n$ sînt unici. În adevăr presupunem prin reducere la absurd că $u = \beta_1 u_1 + \beta_2 u_2 + ... + \beta_n u_n$. Atunci, prin scădere sau egalare, obţinem:

$$(\beta_1 - \alpha_1)u_1 + (\beta_2 - \alpha_2)u_2 + ... + (\beta_n - \alpha_n)u_n = 0$$

Cum vectorii $u_1,...,u_n$ sînt liniar independenţi urmează că $\beta_1 = \alpha_1$, $\beta_2 = \alpha_2$, ... , $\beta_n = \alpha_n$.

Definiţia 2.7.

Scalarii $\alpha_1,\alpha_2,...,\alpha_n$ unic determinaţi senumesc coordonatele vectorului u în baza $u_1,...,u_n$, iar expresia (2) dezvoltarea vectorului u în baza respectivă.

Vrem acum să arătăm că toate bazele din spaţiul S au acelaşi număr de vectori. În acest scop vom demonstra o teoremă cunoscută sub numele de teorema înlocuirii sau teorema lui Steiniz.

Teorema 2.3.

Fie $u_1,...,u_m$ şi $v_1,...,v_n$ vectori din k-spaţiul S, care îndeplinesc condiţiile:

1) $u_1,...,u_m$ liniar independenţi,

2) $u_i \in <v_1,v_2,...,v_n>$ $\quad i = \overline{1,m}$.

În aceste condiţii :

a) $m \leq n$

b) $<v_1,...,v_m,v_{m+1},...,v_n> = <u_1,...,u_m,v_{m+1},...,v_n>$ (eventual după o renumerotare a vectorilor $v_1,...,v_n$).

Demonstraţie: Din $u_1,...,u_m$ liniar independenţi rezultă că $u_i \neq 0 \quad i = \overline{1,m}$, iar din $u_1 \in <v_1,...,v_n> \Rightarrow u_1 = \alpha_1 v_1 + \alpha_2 v_2 + ... + \alpha_n v_n$. Cum $u_1 \neq 0 \Rightarrow$ $(\exists)$ măcar un α nenul. Fie el α_1. Atunci rezultă

(3) $v_1 = -\frac{1}{\alpha_1}u_1 - \frac{\alpha_2}{\alpha_1}v_2 - ... - \frac{\alpha_n}{\alpha_1}v_n$. Relaţia (3) justifică egalitatea:

$$<v_1,v_2,...,v_n> = <u_1,v_2,...,v_n>$$

În continuare reluăm raţionamentul cu vectorul u_2. Din $u_2 \in <v_1,v_2,...,v_n> \Rightarrow u_2 = <u_1,v_2,...,v_n>$ deci:

$$u_2 = \alpha_1 u_1 + \alpha_2 v_2 + ... + \alpha_n v_n$$

În dezvoltarea din partea dreaptă măcar unul dintre $\alpha_2,\alpha_3,...,\alpha_n$ este nenul deoarece în caz contrar ar rămîne $u_2 = \alpha_1 u_1$, relaţie care ar contrazice independenţa vectorilor $u_1,...,u_m$ şi consecinţa 5) a definiţiei 2.5. Presupunem acum $\alpha_2 \neq 0$. Obţinem imediat:

$$v_2 = -\frac{\alpha_1}{\alpha_2}u_1 - \frac{1}{\alpha_2}u_2 - ... - \frac{\alpha_n}{\alpha_2}v_n$$

Această relaţie ne permite să concluzionăm:

$$<u_1,v_2,...,v_n> = <u_1,u_2,v_3,...,v_n>$$

Algoritmul de demonstrare ne arată că vom putea face atîţia paşi cîţi vectori u avem, deci m. Să arătăm că $m>n$ este imposibil.

Dacă $m>n$ atunci după n paşi ajungem la:

$$<v_1,...,v_n> = <u_1,...,u_n>$$

BCU IASI/CENTRAL UNIVERSITY LIBRARY

şi rămîn $u_{n+1}, u_{n+2}, \ldots u_m$ nefolosiţi dar care fac parte din $<u_1, \ldots, u_n>$.

Deci $u_{n+1} \in <u_1, \ldots, u_n> \Rightarrow u_{n+1} = \alpha_1 u_1 + \ldots + \alpha_n u_n$ adică $u_1, \ldots, u_{n+1}$ sînt liniar dependenţi situaţie imposibilă deoarece s-ar contrazice ipoteza. Rămîne numai $m \leq n$ şi atunci şi concluzia b).

Consecinţă: Dacă $u_1, \ldots, u_m$ şi $v_1, \ldots, v_m$ sînt două baze atunci conform teoremei 2.3. rezultă $m \leq n$ şi apoi $n \leq m$ şi deci $m = n$. Aşa dar oricare două baze au acelaşi număr de vectori.

Definiţia 2.8.

Numărul vectorilor dintr-o bază se numeşte dimensiunea spaţiului peste corpul k şi se notează dim S.

Într-un spaţiu liniar de dimensiune finită, orice familie de vectori liniar independenţi poate fi completată pînă la o bază.

Să considerăm acum spaţiul S de dimensiune n şi $u_1, u_2, \ldots, u_n$ o bază.

Admitem că o anume problemă impune folosirea unei noi baze formată din vectori $v_1, v_2, \ldots, v_n$. Dorim să vedem cum se schimbă coordonatele unui vector oarecare x cînd se face trecerea la noua bază.

Fie $x = b_1 u_1 + b_2 u_2 + \ldots + b_n u_n$ şi $x = x_1 v_1 + x_2 v_2 + \ldots + x_n v_n$. În baza $u_1, \ldots, u_n$ vom avea:

$$\begin{cases} v_1 = \alpha_{11} u_1 + \alpha_{12} u_2 + \ldots + \alpha_{1n} u_n \\ v_2 = \alpha_{21} u_1 + \alpha_{22} u_2 + \ldots + \alpha_{2n} u_n \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ v_n = \alpha_{n1} u_1 + \alpha_{n2} u_2 + \ldots + \alpha_{nn} u_n \end{cases}$$

Aceste relaţii determină matricea:

$$A = \begin{pmatrix} \alpha_{11} & \alpha_{21} & \ldots & \alpha_{1n} \\ \ldots\ldots & \ldots\ldots & \ldots\ldots & \ldots\ldots \\ \alpha_{n1} & \alpha_{n2} & \ldots & \alpha_{nn} \end{pmatrix}$$

numită matricea schimbării de bază.

Observaţie: Matricea schimbării de bază este o matrice nedegenerată. Justificarea se realizează astfel: $\det A = 0$ înseamnă că între liniile matricei A există o combinaţie care să dea ultima linie. Deci $l_n = c_1 l_1 + c_2 l_2 + \ldots + c_{n-1} l_{n-1}$. Aceasta însă arată că $v_n = c_1 v_1 + \ldots + c_{n-1} v_n$ adică vectorii $v_1, \ldots, v_n$ ar fi dependenţi, fapt care contrazice ipoteza.

Pentru a găsi formulele de transformare ale coordonatelor înlocuim expresiile vectorilor $v_1, \ldots, v_n$ în dezvoltarea vectorului x. După calcule se obţine:

$$\begin{cases} \alpha_{11}x_1 + \alpha_{21}x_2 + \ldots + \alpha_{n1}x_n = b_1 \\ \alpha_{12}x_1 + \alpha_{22}x_2 + \ldots + \alpha_{n2}x_n = b_2 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \alpha_{1n}x_1 + \alpha_{2n}x_2 + \ldots + \alpha_{nn}x_n = b_n \end{cases}$$

Formulele căutate se obţin după rezolvarea acestui sistem care este un sistem Cramer. Deosebit de important pentru cele ce urmează este cazul în care cele două baze diferă între ele printr-un singur vector.

Să ne ocupăm mai îndeaproape de acest caz. Fie deci $u_1, \ldots, u_n$ baza iniţială şi $u_1, \ldots, u_{n-1}, v$ noua bază şi $x = b_1 u_1 + b_2 u_2 + \ldots + b_n u_n$ şi $x = x_1 u_1 + x_2 u_2 + \ldots + x_{n-1} u_{n-1} + x_n v$. Urmînd o cale analoagă cu cea generală avem:

$$\begin{cases} u_1 = u_1 \\ \ldots\ldots \\ u_{n-1} = u_{n-1} \\ v = \alpha_1 u_1 + \alpha_2 u_2 + \ldots + \alpha_n u_n \end{cases}$$

Matricea schimbării de bază va fi:

$$A = \begin{pmatrix} 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ \alpha_1 & \alpha_2 & \ldots & \alpha_n \end{pmatrix}$$

BCU IAŞI CENTRU DE RESURSE LIBRARY

Din $\det A \neq 0 \Rightarrow \alpha_n \neq 0$.

Deci, în general, vectorul v poate înlocui vectorul u_i din bază dacă coordonata $\alpha_i \neq 0$. Sistemul de ecuaţii va avea forma:

$$\begin{cases} x_1 + \alpha_1 x_n = b_1 \\ x_2 + \alpha_2 x_n = b_2 \\ \dots\dots\dots\dots\dots \\ x_{n-1} + \alpha_{n-1} x_{n-1} = b_{n-1} \\ \alpha_n x_n = b_n \end{cases}$$

De aici obţinem:

$$\begin{cases} x_1 = b_1 - \alpha_1 \dfrac{b_n}{\alpha_n} \\ x_2 = b_2 - \alpha_2 \dfrac{b_n}{\alpha_n} \\ \dots\dots\dots\dots\dots \\ x_{n-1} = b_{n-1} - \alpha_{n-1} \dfrac{b_n}{\alpha_n} \\ x_n = b_n - \alpha_n \dfrac{b_n}{\alpha_n} \end{cases}$$

Formulele găsite permit obţinerea unui algoritm foarte comod de calcul a coordonatelor la o schimbare de bază. În unele cărţi aceste formule sînt numite „regula dreptunghiului". Schimbarea se va face înlocuind numai cîte un vector. Formăm un tabel cu coordonatele tuturor vectorilor consideraţi, notînd atît coordonatele iniţiale cît şi cele finale.

BCU IAŞI CENTRAL UNIVERSITY LIBRARY

Baza	u_1	u_2	.....	u_{n-1}	u_n	v	u
u_1	1	0	.....	0	0	α_1	b_1
u_2	0	1	.....	0	0	α_2	b_2
.	.	.		.	.	.	.
.	.	.		.	.	.	.
.	.	.		.	.	.	.
u_{n-1}	0	0	.....	1	0	α_{n-1}	b_{n-1}
u_n	0	0	.....	0	1	α_n	b_n
u_1	1	0	.....	0	--	0	$b_1 - \alpha_1 \frac{b_n}{\alpha_n}$
u_2	0	1	.....	0	--	0	$b_2 - \alpha_2 \frac{b_n}{\alpha_n}$
.	.	.		.	.	.	.
.	.	.		.	.	.	.
.	.	.		.	.	.	.
u_{n-1}	0	0	.....	1	--	0	$b_{n-1} - \alpha_{n-1} \frac{b}{\alpha}$
v	0	0	.....	0	$\frac{1}{\alpha_n}$	1	$\frac{b_n}{\alpha_n}$

Trecerea de la partea I-a la a-II a se poate realiza prin următoarele două reguli simple şi mecanizabile:

a) Linia vectorului care vine în bază se obţine din linia vectorului care pleacă împărţită la α_n (pivot).

b) Celelalte linii se obţin din vechile linii înmulţite cu cîte un scalar astfel ales încît pe coloana pivotului să se obţină numai de zero (în afara pivotului).

BCU "SIC" CENTRAL UNIVERSITY LIBRARY

Exemplu:

Baza	u_1	u_2	u_3	v	u
u_1	1	0	0	3	5
u_2	0	1	0	-2	1
u_3	0	0	1	4	-3
v	$\frac{1}{3}$	0	0	1	$\frac{5}{3}$
u_2	$\frac{2}{3}$	1	0	0	$\frac{13}{3}$
u_3	$-\frac{4}{3}$	0	1	0	$-\frac{29}{3}$

Am adus în bază în locul lui u_1 vectorul v. În baza v, u_2, u_3 vectorul u se scrie: $u = \frac{5}{3}v + \frac{13}{3}u_2 - \frac{29}{3}u_3$.

Observaţii:

1. Spaţiul R^n are dimensiunea n iar vectorii:

$$e_1 = (1, 0, \ldots 0)$$
$$e_2 = (0, 1, \ldots, 0)$$
$$\ldots\ldots\ldots\ldots$$
$$e_n = (0, 0, \ldots, 1)$$

formează o bază numită bază canonică. Dacă $x \in R^n$ şi $x = (\alpha_1, \alpha_2, \ldots, \alpha_n)$ atunci:

$$x = \alpha_1 e_1 + \alpha_2 e_2 + \ldots + \alpha_n e_n$$

BCU "ASICENTRAL UNIVERSITY LIBRARY

Prin urmare în baza canonică coordonatele vectorului x sînt egale cu componentele de acelaşi nume. Această particularitate este foarte importantă şi folosită în foarte multe cazuri teoretice sau practice.

2. Fie vectorii $u_i = (a_{i1}, a_{i2}, \ldots, a_{in})$ din K^n pentru $i = \overline{1,m}$ şi matricea:

$$A = \begin{pmatrix} u_1 \\ u_2 \\ \ldots\ldots \\ u_m \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots\ldots & \ldots\ldots & \ldots\ldots & \ldots\ldots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix}$$

asociată lor. Sistemul de vectori $u_1, u_2, \ldots, u_m$ este: liniar dependent dacă $rang\, A < m$ şi liniar independent dacă rang $A = m$.

§ 4. OPERATORI LINIARI

Sînt numeroase situaţiile în care avem de a face cu corespondenţe între k-spaţii.

Fie S, T două k-spaţii şi $f: S \to T$

Definiţia 2.9.

Aplicaţia f se numeşte operator liniar (morfism de spaţii sau transformare liniară) dacă

1) $f(u+v) = f(u) + f(v)$

2) $f(\alpha u) = \alpha f(u)$

pentru orice $u, v \in S$ şi orice $\alpha \in k$.

Observaţii:

1) Orice morfism de spaţii este un morfism între grupurile $(S, +)$ şi $(T, +)$. În consecinţă $f(0) = 0$ şi $f(-u) = -f(u)$.

2) Aplicaţia $f: S \to T$ este operator liniar, dacă şi numai dacă $f(\alpha u + \beta v) = \alpha f(u) + \beta f(v)$, $(\forall) u, v \in S$, $(\forall) \alpha, \beta \in k$.

3) Dacă operatorul f este aplicaţie bijectivă se spune că spaţiile S şi T

sînt izomorfe şi se scrie $S \simeq T$.

4) Mulţimea operatorilor liniari $f: S \to T$ se notează cu Hom (S, T) şi este tot un k-spaţiu în raport cu operaţiile $(f+g)(u) = f(u) + g(u)$ şi $(\alpha f)(u) = \alpha f(u)$.

5) Notăm cu:

$$Imf = \{y \in T \mid (\exists) x \in S \text{ cu } f(x) = y\}$$

$$Kerf = \{x \in S \mid f(x) = 0\}$$

Teorema 2.4.

Mulţimile Imf şi $Kerf$ sînt subspaţii liniare în S respectiv în T.

Demonstraţie:

Fie $y_1, y_2 \in Imf \Rightarrow \exists\ x_1, x_2 \in S$ astfel ca $f(x_1) = y_1$, $f(x_2) = y_2$. Putem scrie $y_1 + y_2 = f(x_1) + f(x_2) = f(x_1 + x_2)$ şi cum $x_1 + x_2 \in S \Rightarrow f(x_1 + x_2) = f(x_3) = y_3 \in T$. Pentru $\alpha \in k$ şi $y \in Imf$, $\alpha y = \alpha f(x) = f(\alpha x) \in T$. Deci Imf este subspaţiu în T.

Să considerăm acum mulţimea $Kerf$. Fie $x_1, x_2 \in Kerf$ deci $f(x_1) = f(x_2) = 0$. Deci avem $x_1 + x_2 = x_3$ implică $f(x_3) = f(x_1 + x_2) = f(x_1) + f(x_2) = 0 + 0 = 0 \Rightarrow x_3 \in Kerf$. Apoi dacă $\alpha \in k$ şi $x \in Kerf \Rightarrow f(\alpha x) = \alpha f(x) = \alpha 0 = 0$ adică $\alpha x \in Kerf$. Şi deci afirmaţie este demonstrată.

Observaţie: Spaţiile acestea au importanţă deosebită şi se numesc spaţiul imagine şi nucleul lui f.

Să admitem acum că $\dim S = n$ şi $\dim T = m$.

Teorema 2.5.

Există o corespondenţă bijectivă între mulţimea $Hom(S, T)$ şi $M_{mxn}(k)$ pentru fiecare din bazele alese.

Demonstraţie: Fie $u_1, u_2, \ldots, u_n$ o bază în S şi $v_1, v_2, \ldots, v_m$ o bază din T şi $f \in Hom(S, T)$. Vom avea:

$$\begin{cases} f(u_1) = \alpha_{11}v_1 + \alpha_{12}v_2 + \ldots + \alpha_{1m}v_m \\ f(u_2) = \alpha_{21}v_1 + \alpha_{22}v_2 + \ldots + \alpha_{2m}v_m \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ f(u_n) = \alpha_{n1}v_1 + \alpha_{n2}v_2 + \ldots + \alpha_{nm}v_m \end{cases}$$

Relaţiile date conduc la matricea:

$$A_f = \begin{pmatrix} \alpha_{11} & \ldots & \alpha_{1m} \\ \alpha_{21} & \ldots & \alpha_{2m} \\ \ldots\ldots & \ldots & \ldots\ldots \\ \alpha_{n1} & \ldots & \alpha_{nm} \end{pmatrix}$$

Matricea A_f se numeşte matricea operatorului f în bazele date.

Este valabilă şi reciproca. Fixate bazele şi o matrice A de tip nxm putem defini aplicaţia $f: S \to T$ dînd $f(u_1) = \sum_{j=1}^{m} \alpha_{1j}v_j$, ... , $f(u_n) = \sum_{j=1}^{m} \alpha_{nj}v_j$. Se arată uşor că f este operator liniar şi că $A_f = A$. Corespondenţa aceasta este evident bijectivă şi deci teorema este justificată.

Observaţie: Dacă L,M notează matrici ale schimbărilor de bază din spaţiile S şi T atunci matricea B_f în noua pereche de baze este $B_f = LA_fM^{-1}$.

Definiţia 2.10.

Spaţiile S şi T se numesc izomorfe dacă între ele există un operator bijectiv.

Teorema 2.6.

Orice spaţiu liniar S peste corpul k de dimensiune n este izomorf cu spaţiul k^n.

Demonstraţie:

Fie $u_1,u_2,\ldots,u_n$ o bază în S şi $x \in S$. Atunci avem scrierea unică $x = \alpha_1u_1 + \alpha_2u_2 + \ldots + \alpha_nu_n$.

Să considerăm aplicaţia $x \to f(x) = (\alpha_1, \alpha_2, ..., \alpha_n)$. Se arată cu uşurinţă că această aplicaţie este bijectivă şi morfism de spaţii.

Consecinţă: Toate spaţiile vectoriale de aceeaşi dimensiune peste acelaşi corp de bază sînt izomorfe. Din acest motiv toate aplicaţiile numerice se fac de regulă pe spaţiile $\mathbb{R}^n$, $n \in \mathbb{N}^*$.

§ 5. FORME LINIARE, BILINIARE ŞI PĂTRATICE

Să ne ocupăm acum de cazul particular cînd spaţiul T este chiar k.

Definiţia 2.11.

Un operator liniar $f: S \to k$ se numeşte formă liniară. Dacă în S fixăm baza $u_1, u_2, ..., u_n$ şi $x = x_1u_1 + x_2u_2 + ... + x_nu_n$ atunci $f(x) = x_1f(u_1) + ... + x_nf(u_n)$ şi dacă notăm $\alpha_i = f(u_i)$ urmează că $f(x) = \alpha_1x_1 + ... + \alpha_nx_n$. Deci într-o bază fixată orice formă liniară determină un polinom liniar şi omogen în coordonatele vectorului. Este valabilă şi afirmaţia reciprocă.

Conform observaţiei 4 de la definiţia 2.9. rezultă ca $Hom(S, k)$ este un k-spaţiu liniar. Acest spaţiu se notează de obicei cu S^* şi se numeşte spaţiul dual lui S. Mai mult putem demonstra şi că $dim\, S^* = dim\, S = n$. De la forme liniare necesităţile practice au impus generalizări ale lor.

Să reluăm S, T două $\mathbf{R}$-spaţii.

Definiţia 2.12.

O aplicaţie $f: S \times T \to \mathbb{R}$ se numeşte formă biliniară dacă are proprietăţile:

1) $f(x_1 + x_2, y) = f(x_1, y) + f(x_2, y)$ 2) $f(\alpha x, y) = \alpha f(x, y)$

3) $f(x, y_1 + y_2) = f(x, y_1) + f(x, y_2)$ 4) $f(x, \beta y) = \beta f(x, y)$

pentru orice $x \in S$, $y \in T$, $\alpha \in \mathbb{R}$.

Se poate spune că forma biliniară este o formă liniară de fiecare argument. Această observaţie simplă este foarte utilă la trecerea la forme triliniare apoi la forme multiliniare.

Teorema 2.7.

Oricărei forme biliniare, într-o pereche de baze, îi corespunde o matrice şi reciproc.

Demonstraţie: Fie $\dim S = n$, $\dim T = m$, $u_1, u_2, \ldots, u_n$ şi respectiv $v_1, v_2, \ldots, v_m$ o pereche de baze în cele două $\mathbb{R}$-spaţii. Pentru $x \in S$, $y \in T$ şi $x = x_1u_1 + \ldots + x_nu_n$, $y = y_1v_1 + \ldots + y_mv_m$ putem scrie:

$$f(x,y) = f(x_1u_1 + \ldots + x_nu_n, y_1v_1 + \ldots + y_mv_m) = \sum_{i=1}^{n}\sum_{i=1}^{m} f(u_i, v_j)x_iy_j$$

Dacă notăm $\alpha_{ij} = f(u_i, v_j) \Rightarrow f(x,y) = \sum_{i=1}^{n}\sum_{j=1}^{m} \alpha_{ij}x_iy_j$. Matricea $A = (\alpha_{ij})$ unde $\alpha_{ij} = f(u_i, v_j)$ se numeşte matricea formei biliniare în bazele fixate iar elementele α_{ij} coeficienţii formei date. Demonstraţia reciprocei o lăsăm cititorului. Dacă în S facem o schimbare de bază de matrice L iar în T o schimbare de bază de matrice M, atunci noua matrice a formei va fi $B = LAM^T$ (M^T transpusa lui M). În cazul particular în care spaţiile S şi T coincid noua matrice are forma $B = LAL^T$.

Să notăm cu $S_0 = \{x \in S \mid f(x,y) = 0 \ (\forall) y \in T\}$ şi

$$T_0 = \{y \in T \mid f(x,y) = 0 \quad (\forall) x \in S\}$$

Se verifică cu uşurinţă că mulţimile S_0 şi T_0 sînt subspaţii liniare în S respectiv T. Aceste subspaţii liniare se numesc subspaţii nule pentru f.

Dacă $S_0 = 0$ şi $T_0 = 0$ forma f se numeşte formă biliniară nedegenerată.

Mai facem şi observaţia deosebit de utilă că o alegere convenabilă de baze în spaţiile S, T în care primii vectori să fie din S_0 respectiv T_0 are ca efect obţinerea matricii noi de forma (r - numindu-se rangul formei biliniare):

$$\begin{pmatrix} \alpha_{11} \ldots \alpha_{1r} & 0 \ldots 0 \\ \ldots\ldots\ldots\ldots & \\ \alpha_{r1} \ldots \alpha_{rr} & 0 \ldots 0 \\ 0 \ldots\ldots 0 & 0 \ldots\ldots 0 \\ \ldots\ldots\ldots\ldots & \\ 0 \ldots\ldots 0 & 0 \ldots\ldots 0 \end{pmatrix}$$

De fapt există posibilitatea să alegem bazele în aşa fel încît matricea să fie chiar de forma diagonală. Forma de prezentare a formei biliniare, în acest caz, se numeşte forma canonică. Aşa dar forma canonică este:

$$f(x,y) = x_1y_1 + x_2y_2 + \ldots + x_ry_r$$

Numeroase sînt situaţiile în care $S = T$. În aceste cazuri matricea formei este pătratică de ordin n iar forma se zice definită pe $S \times S$ (sau pe S).

Definiţia 2.13.

O formă biliniară definită pe S, unde S este $\mathbb{R}$ spaţiu este simetrică dacă:

$$f(x,y) = f(y,x)$$

şi antisimetrică dacă $f(x,y) = -f(y,x)$ $(\forall)$ $x,y \in S$.

Observaţie: O formă biliniară simetrică are matricea simetrică iar o formă antisimetrică are matricea antisimetrică în orice bază.

Importanţa acestor tipuri de forme este ilustrată şi de afirmaţia: orice formă biliniară definită pe spaţiul S este sumă de două forme: una simetrică şi alta antisimetrică.

Definiţia 2.14.

Dacă f este o formă biliniară simetrică definită pe $\mathbb{R}$-spaţiul S se numeşte formă pătratică funcţia $p : S \to \mathbb{R}$ dată de:

$$p(x) = f(x,x)$$

Între formele biliniare simetrice şi formele pătratice există o corespondenţă bijectivă.

Dată f-formă biliniară simetrică avem determinată funcţia p.

Reciproc. Dată forma pătratică p să notăm cu:

$$f(x,y) = \frac{1}{2}(p(x+y) - p(x) - p(y))$$

Rezultă imediat că $f(x,y)$ reprezintă o formă biliniară definită pe S.

Această funcţie este cunoscută sub denumirea de forma biliniară polară a formei pătratice p.

Definiţia 2.15.

Forma biliniară simetrică $f(x,y)$ este pozitiv (negativ) definită dacă

$$p(x) > 0 \ (\forall)\ x \in S \quad x \neq 0$$

respectiv:

$$p(x) < 0 \ (\forall)\ x \in S \ x \neq 0$$

Legătura strînsă între formele pătratice şi formele biliniare simetrice ne ajută să transpunem proprietăţile de la formele biliniare la cele pătratice. Aşa de exemplu se obişnuieşte să se vorbească de rangul formei pătratice care este, cum este natural, rangul matricii formei. Acest rang este invariant la transformări de baze. Mai mult şi formele pătratice au forme canonice. Orice formă pătratică de rang r se poate aduce la o formă canonică prin transformări de baze (sau de coordonate) convenabile.

Oricum am proceda pentru a obţine o formă canonică numărul pătratelor pozitive (deci şi numărul pătratelor negative) este acelaşi. Aceasta este "legea de inerţie" a lui Sylvester.

Metoda lui Gauss (metoda grupărilor) de aducere la forma canonică.

În această metodă transformările se fac asupra coordonatelor. Notăm

$$p(x) = F(x_1, x_2, \ldots, x_n) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_i y_j \text{ cu } a_{ij} = a_{ji}.$$

Presupunem $a_{11} \neq 0$. Grupăm toţi termenii care conţin pe x_1 şi formăm cu ei pătratul unei forme liniare.

$$F(x_1,x_2,\ldots,x_n) = (a_{11}x_1^2 + 2a_{12}x_1x_2 + 2a_{13}x_1x_3 + \ldots + 2a_{1n}x_1x_n) +$$

$$+ \sum_{i=2}^{n}\sum_{j=2}^{n} a_{ij}x_ix_j = \frac{1}{a_{11}}\left[\sum_{j=1}^{n} a_{1j}x_j\right]^2 - \sum_{j=2}^{n}(a_{1j}x_j)^2 + \sum_{i=2}^{n}\sum_{j=2}^{n} a_{ij}x_ix_j$$

Dacă notăm paranteza pătrată cu y_1 se obţine:

$$F(x_1,x_2,...,x_n)=\frac{1}{a_{11}}y_1^2+F_1(x_2,...,x_n)$$

unde F_1 este tot o formă pătratică cu $n-1$ nedeterminate. Apoi procedeul se repetă pînă se ajunge la forma canonică.

Dacă $a_{11}=0$ dar $a_{pp}\neq 0$ facem $x_1=y_p$ şi $x_p=y_1$ şi restul $x_i=x_i$ şi obţinem $a'_{11}\neq 0$.

Dacă toţi $a_{ii}=0$ atunci există măcar un $a_{ij}\neq 0$ şi făcînd transformarea $x_i=z_i+z_j$, $x_j=z_i-z_j$, $x_k=z_k$ pentru $k\neq i,j$. Se ajunge la $a_{ii}\neq 0$ apoi poate fi făcut a_{11}. Acum putem enunţa:

Teorema 2.8.

Condiţia necesară şi suficientă ca o formă pătratică să fie pozitiv definită este ca în forma canonică să nu avem decît pătrate cu coeficienţi pozitivi.

§ 6. SPAŢII EUCLIDIENE

Să considerăm acum un spaţiu liniar E peste $\mathbb{R}$.

Definiţia 2.16.

Spaţiul liniar E se numeşte spaţiu euclidian dacă între elementele sale este introdus produsul scalar notat (x,y) prin care perechii x,y i se asociază un număr real cu proprietăţile:

1) $(x,y)=(y,x)\quad(\forall)\,x,y\in E$,

2) $(x_1+x_2,y)=(x_1+y)+(x_2+y)\quad(\forall)\,x_1,x_2,y\in E$,

3) $(\alpha x,y)=\alpha(x,y)\quad(\forall)\,\alpha\in\mathbb{R}$, $\quad(\forall)\,x,y\in E$,

4) $(x,x)\geq 0$ şi $(x,x)=0$ implică $x=0$, $(\forall)\,x\in E$.

Consecinţă $(x,y_1+y_2)=(x,y_1)+(x,y_2)$ şi $(x,\alpha y)=\alpha(x,y)$.

Demonstraţia este imediată.

Exemplu:

Spaţiul $\mathbb{R}^n$ devine euclidian dacă definim:

$$(x,y) = x_1y_1 + x_2y_2 + \ldots + x_ny_n$$

Acest spaţiu euclidian se mai notează cu E_n.

Fie E un spaţiu euclidian.

Definiţia 2.17.

a) Expresia $\sqrt{(x,x)} = |x|$ se numeşte lungimea vectorului x.

b) Doi vectori $x,y \in E$ se numesc ortogonali dacă produsul scalar al lor este zero: $(x,y) = 0$.

c) Baza $e_1,\ldots,e_n$ se numeşte ortonormat dacă $(e_i,e_j) = \begin{cases} 1 & i=j \\ 0 & i \neq j \end{cases}$

Teorema 2.9.

Dacă $e_1,\ldots,e_n$ sînt vectori nenuli care satisfac condiţiile $(e_i,e_j) = 0$ dacă $i \neq j$, atunci vectorii $e_1,\ldots,e_n$ sînt liniar independenţi.

Demonstraţie: Presupunem că are loc relaţia:

$$\alpha_1e_1 + \alpha_2e_2 + \ldots + \alpha_ke_k = 0$$

Mai avem şi $(e_i, \alpha_1e_1 + \alpha_2e_2 + \ldots + \alpha_ke_k) = 0$.

De aici rezultă $\alpha_i(e_i,e_i) = 0$ şi de aici $\alpha_i = 0 \quad i=\overline{1,k}$.

Teorema 2.10. (Metoda lui Schmidt)

În orice spaţiu euclidian cu n dimensiuni există baze ortonormate.

Demonstraţie: Fie $f_1,f_2,\ldots,f_n$ o bază oarecare din spaţiul E. Facem următoarea schimbare de bază:

$$e_1 = f_1$$
$$e_2 = f_2 - \alpha_2{}^1 f_1$$
$$e_3 = f_3 - \alpha_3{}^1 f_1 - \alpha_3{}^2 f_2$$
$$\ldots\ldots\ldots\ldots\ldots$$

BCU IAŞI BIBLIOTECA CENTRALĂ UNIVERSITARĂ

Determinăm numerele α_i^j astfel încît să fie îndeplinite condiţiile $(e_i,e_j) = 0 \quad i \neq j$.

Din $(e_2,e_1) = (e_2,f_1) = (f_2 - \alpha_2^1 f_1, f_1) = 0$ obţinem:

$$\alpha_2^1 = \frac{(f_2,f_1)}{(f_1,f_1)}, \quad (f_1,f_1) \neq 0$$

Presupunem inductiv construiţi vectorii $e_1, e_2, \ldots, e_k$ ortogonali doi cîte doi. Pentru vectorul e_{k+1} impunem condiţiile:

$$(e_{k+1}, e_h) = 0 \qquad h = 1,2,\ldots,k$$

Dînd lui h valorile $1,2,\ldots,k$ se obţine un sistem de k ecuaţii în necunocutele $\alpha_{k+1}^1, \alpha_{k+1}^2, \ldots, \alpha_{k+1}^k$.

Din $(e_h,e_j) = 0$ pentru $h \neq j$ şi $h,j \leq k$ obţinem că fiecare ecuaţie capătă forma:

$$-\alpha_{k+1}^h (e_h,e_h) + (e_h,f_{k+1}) = 0 \qquad h = 1,2,\ldots,k$$

De aici scot α_{k+1}^2. Vectorii e_i obţinuţi sînt ortogonali. În continuare normarea se obţine făcînd schimbarea:

$$e_i^* = \frac{e_i}{\sqrt{(e_i,e_i)}}, \quad i = \overline{1,n} \quad \text{avem } (e_i^*,e_i^*) = 1 \quad i = \overline{1,n}.$$

§ 7. SISTEME DE ECUAŢII LINIARE

Scopul acestui paragraf este de a completa cunoştinţele despre sisteme căpătate în liceu. Datorită acestei intenţii nu am mai reluat chestiunile învăţate, în clasa a XI-a de liceu.

A) Sisteme descrise vectorial

Aşa cum ştim forma generală a unui sistem de ecuaţii liniare este:

$$(1)\ \begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_n \end{cases} \qquad \alpha_{ij} \in \mathbb{R}.$$

Să convenim să notăm cu A - matricea sistemului, B - matricea extinsă, $P_1, P_2, \ldots, P_n$ coloanele matricii A şi P_0 coloana termenilor liberi.

Sistemul considerat se scrie evident:

(2) $AX = P_0$

sau (3) $x_1P_1 + x_2P_2 + \ldots + x_nP_n = P_0$.

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{pmatrix}$$

Dacă $P_0 = 0$ sistemul se numeşte omogen.

Definiţia 2.18.

a) Se numeşte soluţie a sistemului de ecuaţii orice vector $X \in \mathbb{R}^n$ care verifică sistemul.

b) Un sistem care are măcar o soluţie se numeşte compatibil.

Lemă. Dacă un sistem are două soluţii X_1, X_2 atunci are o infinitate de soluţii.

Demonstraţie:

Avem $AX_1 = P_0$, $AX_2 = P_0$. Fie $Y = \lambda X_1 + (1-\lambda)X_2$, $\lambda \in \mathbb{R}$. Atunci putem scrie:

$$AY = A[\lambda X_1 + (1-\lambda)X_2] = \lambda AX_1 + (1-\lambda)AX_2 = \lambda P_0 + (1-\lambda)P_0 = P_0$$

Deci Y este tot soluţie.

Această observaţie permite clasificarea sistemelor compatibile în două tipuri:

BCU "MIRCEA CEL BĂTRÂN" UNIVERSITY LIBRARY

a) Sisteme compatibile cu soluţie unică - sisteme determinate.

b) sistem compatibil cu măcar două soluţii numit compatibil nedeterminat.

Teorema 2.11. (Teorema generală de compatibilitate)

Condiţia necesară şi suficientă ca sistemul (3) să fie compatibil este ca $P_0 \in <P_1, P_2, ..., P_n>$.

Demonstraţie: Condiţia este necesară. Dacă sistemul este compatibil şi $X=(\alpha, \alpha_2, ..., \alpha_n)$ este o soluţie putem scrie:

$P_0 = \alpha_1 P_1 + \alpha_2 P_2 + ... + \alpha_n P_n$ adică $P_0 \in <P_1, P_2, ..., P_n>$.

Condiţia este suficientă: Dacă $P_0 \in <P_1, P_2, ..., P_n> \rightarrow$ există $\alpha_1, \alpha_2, ..., \alpha_n$ aşa încît $\alpha_1 P_1 + ... + \alpha_n P_n = P_0$ adică $X = (\alpha_1, \alpha_2, ..., \alpha_n)$ este o soluţie.

Observaţie: Teorema 2.11 este echivalentă cu teorema lui Kronecker-Capelli sau teorema Rouché întîlnite în manualele de liceu.

Să observăm că de fiecare sistem se poate ataşa cu uşurinţă un sistem omogen ignorînd termenii liberi. Se obţin:

$$(1^1)\begin{cases} a_{11}x_1 + a_{12}x_2 + ... = a_{1n}x_n = 0 \\ a_{21}x_1 + a_{22}x_2 + ... + a_{2n}x_n = 0 \\ ..... \\ a_{m1}x_1 + a_{m2}x_2 + ... + a_{mn}x_n = 0 \end{cases}$$

$$(2^1)\ AX = 0$$

$$(3^1)\ x_1 P_1 + x_2 P_2 + ... + x_n P_n = 0$$

Teorema 2.12.

Orice sistem omogen este compatibil şi are şi soluţii cu componente nenule dacă şi numai dacă $r_A < n$.

Demonstraţie:

Pentru prima parte este suficient să observăm că $x_1 = x_2 = ... = x_n = 0$ este soluţie. Pentru a fi nedeterminat trebuie ca măcar o necunoscută să fie

secundară deci rangul matricei A să fie mai mic decît n.

Teorema 2.13.

Mulţimea soluţiilor unui sistem omogen formează un subspaţiu liniar în R^n de dimensiune $n - r$ unde $r = r_A$.

Demonstraţie: Fie X_1 şi X_2 două soluţii. Atunci $X_1 + X_2$ şi αX_1 cu $\alpha \in R$ sînt tot soluţii. În adevăr $A(X_1 + X_2) = AX_1 + AX_2 = 0 + 0 = 0$ şi $A(\alpha X_1) = \alpha(AX_1) = \alpha 0 = 0$.

Să găsim dimensiunea şi o bază în acest spaţiu liniar de soluţii.

Admitem că $r_A = r$ şi că:

$$\delta = \begin{vmatrix} a_{11} & \dots & a_{1r} \\ \dots & \dots & \dots \\ a_{r1} & \dots & a_{rr} \end{vmatrix}$$

este determinantul principal. Atunci pentru rezolvare vom opri numai ecuaţiile principale pe care le vom scrie în forma:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1r}x_r = -a_{1,r+1}x_{r+1} - \dots - a_{1n}x_n \\ a_{21}x_2 + a_{22}x_2 + \dots + a_{rr}x_r = -a_{2,r+1}x_{r+1} - \dots - a_{2n}x_n \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ a_{r1}x_1 + a_{r2}x_2 + \dots + a_{rr}x_r = -a_{r,r+1}x_{r+1} - \dots - a_{rn}x_n \end{cases}$$

Rezolvînd acest sistem, în ipoteza $x_{r+1}, \dots, x_n$ cunoscute vom obţine nişte expresii de forma:

$$\begin{cases} x_1 = b_{11}x_{r+1} + b_{12}x_{r+2} + \dots + b_{2,r-n}x_n \\ x_2 = b_{21}x_{r+1} + b_{22}x_{r+2} + \dots + b_{2,r-n}x_n \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ x_r = b_{r1}x_{r+1} + b_{r2}x_{r+2} + \dots + b_{r,r-n}x_n \end{cases}$$

Pentru a obţine diverse soluţii particulare vom da necunoscutelor secundare valori după voie. Să facem pe rînd cîte una din ele egale cu 1 şi restul zero vom obţine:

BCU IASI / CENTRAL UNIVERSITY LIBRARY

x_1	x_2	...	x_r	x_{r+1}	x_{r+2}	...	x_n
b_{11}	b_{21}	...	b_{r1}	1	0	...	0
b_{21}	b_{22}	...	b_{r2}	0	1	...	0
...	...	...	...	...	...	...	...
$b_{1,n-r}$	$b_{2,n-r}$	...	$b_{r,n-r}$	0	0	...	1

Se obţin $n - r$ soluţii pe care le notăm cu $X_1, X_2, ...X_{n-r}$ avînd componentele înregistrate pe linie.

După cum se poate verifica imediat aceşti vectori sînt liniar independenţi şi dacă $X = (a_1, a_2, ..., a_n)$ este o soluţie oarecare atunci $X = a_{r+1}X_1 + a_{r+2}X_2 + ... + a_nX_{n-r}$. Prin urmare vectorii $X_1, X_2, ...X_{n-r}$ generează spaţiul soluţiilor şi chiar mai mult ei formează o bază deoarece sînt şi independenţi. O astfel de bază este numită, prin tradiţie, sistem fundamental de soluţii.

Deoarece problema dependenţei sau independenţei liniare a unui sistem de m vectori din R^n este strîns legată de existenţa soluţiilor nebanale ale unui sistem omogen putem concluziona: dacă $m > n$ vectorii sînt liniari dependenţi, dacă $m \leq n$ atunci vectorii sînt independenţi dacă şi numai dacă rangul matricii componentelor lor este egal cu numărul vectorilor.

Metoda eliminării complete

Metoda lui Gauss (a „eliminării complete") pentru rezolvarea sistemelor liniare este dintre cele mai avantajoase din punct de vedere practic. Să considerăm sistemul (S) alcătuit din m ecuaţii liniare

$$(E_i) \qquad a_{i1}x_1 + a_{i2}x_2 + ... + a_{in}x_n = b_i$$

în necunoscutele $x_1, x_2, ..., x_n$.

Renumerotînd eventual ecuaţiile (E_i) putem admite că are loc $a_{11} \neq 0$.

Amplificînd succesiv (E_1) cu factori convenabil aleşi şi scăzînd din ecuaţiile (E_i) cu $i \neq 1$ ne putem asigura că x_1 nu mai figurează în celelalte ecuaţii. Deşi coeficienţii iniţiali s-au modificat vom păstra aceleaşi notaţii ca la început.

Renumerotînd eventual ecuaţiile (E_i) cu $i = \overline{2,m}$ putem admite că are loc $a_{22} \neq 0$. Amplificînd succesiv (E_2) cu factori convenabil aleşi putem face să dispară necunoscuta x_2 atît din ecuaţiile (E_i) cu $i > 2$ dar şi din ecuaţia (E_1). (Ultima operaţie reluată şi pentru variabile $x_3,\ldots,x_r$ justifică adjectivul „complet" din denumirea metodei).

Să admitem că am reluat ultima etapă şi pentru $x_3,\ldots,x_{k-1}$. În încercarea de a relua operaţia pentru variabila x_k putem constata că în ecuaţiile $(E_k),(E_{k+1}),\ldots,(E_m)$ nu mai figurează variabila x_k. Altfel spus, pentru coeficienţii actualizaţi are loc:

$$(*)\ a_{ik} = 0 \quad \text{pentru} \quad i \geq k.$$

În această situaţie vom face operaţia mai sus descrisă la variabila x_2 pentru variabila x_h cu $h > k$, desigur în ipoteza că ecuaţiile (E_i) cu $i = \overline{k,m}$ mai conţin măcar o variabilă x_h. Operaţia va fi încheiată cînd:

- nu mai avem variabile x_h de luat în consideraţie, sau
- nu mai avem ecuaţii care să permită eliminarea unei noi variabile.

Putem rezuma rezultatul acestor operaţii de „reducere" sau „eliminare" considerînd că după eventuale renumerotări ale necunoscutelor x_i prin y_j şi renotări de coeficienţi sistemul se prezintă sub forma:

$$\begin{aligned}
(E_1)\quad & d_1y_1 + \quad + c_{11}y_{r+1} + \ldots + c_{1p}y_n = e_1\\
(E_2)\quad & \quad d_2y_2 + \quad + c_{21}y_{r+1} + \ldots + c_{2p}y_n = e_2\\
& \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
& \quad d_ry_r + c_{rp}y_{r+1} + \ldots + c_{rp}y_n = e_r\\
& \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ 0 = f_{r+1}\\
& \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
(E_m)\quad & \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ 0 = f_m
\end{aligned}$$

Concluziile sînt acum imediate:

(α) Dacă există cel puţin un indice j încît $f_j \neq 0$ sistemul este incompatibil.

(β) Dacă $f_{r+1} = f_{r+2} = \ldots = f_m = 0$ sistemul este compatibil cu subcazurile:

(β_1) Dacă $r = n$ soluţia este unică, dată de $y_i = e_i : d_i$ pentru $i = \overline{1,r}$.

(β_2) Dacă $r < n$, se pot da valori arbitrare variabilelor y_j cu $j > r$, celelalte fiind unic precizate de ecuaţiile (E_i) cu $i = \overline{1,r}$.

Succesiunea operaţiilor mai sus descrise se realizează grafic simplu şi clar folosind în locul ecuaţiilor tablouri de coeficienţi. Spre a preîntîmpina confuzii tablourile conţin o primă linie cu indicele variabilei corespunzătoare. Vom prezenta o astfel de succesiune de tablouri pentru a rezolva efectiv următoarea problemă.

Să se precizeze pentru ce valori ale parametrului real m sistemul:

$$\begin{cases} x_1 + 3x_2 + 4x_3 + 2x_4 - 3x_5 + 5x_6 = 13 \\ x_1 + 6x_2 + 5x_3 + \quad + x_5 + 6x_6 = 16 \\ -x_1 + 3x_2 - 2x_3 - 8x_4 + 14x_5 - 4x_6 = -4 \\ x_1 + 9x_2 + 6x_3 + 2x_4 - x_5 + 9x_6 = 13 + 2m \end{cases}$$

este compatibil şi, pentru aceste valori ale lui m să i se indice soluţia generală. Tablourile sînt în succesiunea lor:

(T1)

1	2	3	4	5	6	
1	3	4	2	-3	5	13
1	6	5	0	1	6	16
-1	3	-2	-8	14	-4	-4
1	9	6	2	-1	9	13+2m

(T2)

1	2	3	4	5	6	
1	3	4	2	-3	5	13
0	3	1	-2	4	1	3
0	6	2	-6	11	1	9
0	6	2	0	2	4	2m

(T3)

1	2	3	4	5	6	
1	0	3	4	-7	4	10
0	3	1	-2	4	1	3
0	0	0	-2	3	-1	3
0	0	0	2	-3	1	m-3

(T4)

1	2	3	4	5	6	
1	0	0	3	-1	2	16
0	3	0	1	1	2	0
0	0	-2	0	3	-1	3
0	0	0	0	0	0	m

Concluzii . Sistemul este compatibil dacă (şi numai dacă) $m = 0$. Pentru soluţia generală x_3, x_5, x_6 pot lua valori arbitrare şi:

$$x_1 = 16 - 3x_3 + x_5 - 2x_6$$

$$x_2 = -\frac{1}{3}(x_3 + x_5 + 2x_6)$$

$$x_4 = \frac{1}{2}(3x_5 - x_6 - 3)$$

§ 8. POLINOAME DE INTERPOLARE

Numeroase aplicaţii ale matematicii în economie includ următoarea problemă.

Sînt date valori numerice x_i, y_i cu $i = \overline{0,n}$ şi astfel încît $i \neq j \Rightarrow x_i \neq x_j$. Se cere să se determine polinomul P de grad minim ce satisface condiţiile:

BCU IAŞI/CENTRUL UNIVERSITAR LIBRARY

$(L)\ P(x_i) = y_i,\ i = \overline{0,n}.$

(Marcarea prin (L) a acestor condiţii evocă numele matematicianului Lagrange ce a pus, a rezolvat şi a utilizat această problemă).

Deoarece în (L) figurează $n+1$ condiţii vom căuta P având $n+1$ coeficienţi deci de grad n, adică

$(P)\ P(x) = a_0x^n + a_1x^{n-1} + \ldots + a_{n-1}x + a_n.$

Astfel sistemul de condiţii (L) induce un sistem liniar de $n+1$ ecuaţii în cele $n+1$ necunoscute: $a_0, a_1, \ldots, a_n$.

Determinantul sistemului liniar este alcătuit din linii.

$$V_i = |x_i^n \quad x_i^{n-1} \ \ldots \ x_i \quad 1|, i = \overline{0,n}$$

se numeşte „determinant Vandermondé" şi notăm valoarea sa prin $V_{\overline{0,n}}(x_i)$.

Scăzînd linia ultimă din primele n şi dezvoltînd după ultima coloană şi scoţînd factori pe linii obţinem o egalitate:

$$(V)\quad V_{\overline{0,n}}(x_i) = [\prod_{i=0}^{n-1}(x_i - x_n)] \cdot V^1,$$

determinantul V^1 va fi adus rapid la forma $V^1 = V_{\overline{0,n}}(x_i)$. Astfel egalitatea (V) devine formulă de recurenţă şi asigură:

$$(W)\qquad V_{\overline{0,n}}(x_i) = \prod_{0 \le i < j \le n}(x_i - x_j).$$

Fapt esenţial pentru consideraţiile noastre este că are loc $V_{\overline{0,n}}(x_i) \neq 0$, deci (L) este sistem Cramer ce va permite determinarea unică a necunoscutelor a_j cu $j = \overline{0,n}$. Semnalăm acum că pentru date concrete (x_i, y_i), $i = \overline{0,n}$ este posibil să rezulte $a_0 = 0$ deci polinomul P are gradul cel mult egal cu n. Pentru a exprima polinomul P sub formă de determinant vom considera şi „necunoscuta" $a_{n+1} = -1$ şi vom transcrie egalităţile (P) şi (L) sub forma:

$$\begin{cases} a_0x^n + a_1x^{n-1} + \ldots + a_{n-1}x + a_n + P(x)a_{n+1} = 0 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_0x_i^n + a_1x_i^{n-1} + \ldots + a_{n-1}x_i + a_n + y_ia_{n+1} = 0 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_0x_n^n + a_1x_n^{n-1} + \ldots + a_{n-1}x_n + a_n + y_na_{n+1} = 0 \end{cases}$$

sistem liniar şi omogen în necunoscutele a_i cu $i = \overline{0,n+1}$ ce admite soluţie netrivială ($a_{n+1} = -1$) deci are determinantul nul. Rezultă astfel formula lui Lagrange sub formă de determinant:

$$(L,D) \quad P(x)V_{\overline{0,n}}(x_i) + W(x,x_i,y_i) = 0$$

unde s-a notat:

$$W(x,x_i,y_i) = \begin{vmatrix} x^n & x^{n-1} & \dots & 1 & 0 \\ x_0^n & x_0^{n-1} & \dots & 1 & y_0 \\ \dots & \dots & \dots & \dots & \dots \\ x_n^n & x_n^{n-1} & \dots & 1 & y_n \end{vmatrix}$$

Considerînd modul în care intervin valorile y_i , $i=\overline{0,n}$ în soluţia Cramer a sistemului L vom constata că vor exista polinoame $Q_i(x)$, fiecare de grad n, independente de valorile y încît:

$$(Q) \quad P(x) = y_0Q_0(x) + y_1Q_1(x) + \dots + y_nQ_n(x).$$

(Polinoamele Q_j vor constitui o bază în spaţiul liniar S al polinoamelor de grad cel mult n şi $\mathcal{P}$ va avea coordonatele y_j în această bază).

Pentru un astfel de polinom Q_j deducem

$$i \neq j \qquad Q_j(x_i) = 0 \qquad Q_j(x_j) = 1.$$

Considerăm polinomul de grad $(n+1)$

$$A(x) = (x-x_0)(x-x_1)\dots(x-x_n)$$

şi polinoamele de grad n

$$C_j(x) = \frac{A(x)}{x-x_j}$$

Urmează imediat $C_j(x_j) \neq 0$, deci putem considera numere reale q_j încît $q_j \cdot C_j(x_j) = 1$. Deducem acum rapid că singurul polinom Q_j ce satisface condiţiile evidenţiate este:

$$Q_j(x_j) = q_j \cdot C_j(x)$$

Astfel egalitatea (Q) asigură „formula lui Lagrange sub formă de sumă"

$$(LS) \qquad P(x) = \sum_{j=0}^{n} y_j \cdot q_j \cdot C_j(x).$$

Observaţii: Problema se reformulează în condiţii concrete cerînd determinarea valorilor unor parametri a_i încît rolul polinomului $P(x)$ să îl preia o funcţie $F(a_i, x)$ şi induce „metoda celor mai mici pătrate" cînd se consideră că valorile y_i sînt afectate de erori.

Întrebări de autocontrol, exerciţii.

1. Fie $S = \{a + b\sqrt{2} + c\sqrt{3} \mid a, b, c \in Q\}$. Să se arate că S este Q—spaţiu liniar. Daţi exemplu de subspaţii ale lui S.

2. În spaţiul vectorial $k[x]$ precizaţi care dintre următoarele condiţii asupra polinoamelor P conduce la subspaţii: grad $P = n$, grad $P \leq n$, grad $P \geq n$, unde $n \in \mathbb{N}$ şi n este fixat .

3. Care din spaţiile vectoriale din exemplele de mai sus este finit generat?

4. Fie S un $\mathbb{C}$—spaţiu, $\dim S = n$ şi $f: S \to S$ un operator liniar. Dacă $f(v) = \lambda v$ pentru un $v \neq 0$ se spune că v este vector propriu şi λ respectiv valoare proprie.

Arătaţi că există un polinom $P \in C[x]$ cu grad $P = n$ care admite ca rădăcini valorile proprii şi că pentru fiecare valoare proprie a lui f mulţimea vectorilor proprii formează un subspaţiu în S.

5. Se dau vectorii $u_1 = (1,0,3)$, $u_2 = (2,1,-1)$, $u_3 = (0,3,1)$. Precizaţi vectorul $u_4 = 3u_1 + 2u_2 - u_3$. Se pot preciza coordonatele lui u_3 în baza u_1, u_2, u_4?

6. Se dau vectorii: $u_1 = (1,2,3)$, $u_2 = (3,-1,2)$, $u_3 = (0,0,0)$, $u_4 = (4,1,5)$, $u_5 = (5,3,8)$, $u_6 = (6,5,11)$.

Să se aleagă dintre ei un număr maxim de vectori independenţi şi să se exprime ceilalţi drept combinaţii liniare de cei aleşi.

7. Fie $A = X^2 - 3X + 1$, $B = X^2 + 5X - 3$, $C = X^2 + X - 1$ elemente din $\mathbf{R}[X]$. Să se spună dacă sînt sau nu liniar independente.

8. Preluînd polinoamele A, B, C din exerciţiul precedent, fie $D = A+\sqrt{2}B$, $E = B+\sqrt{2}C$, $F = C+\sqrt{2}A$. Să se spună dacă D, E, F, ca elemente ale spaţiului vectorial $K[X]$ sînt sau nu independente cînd: a) K este $\mathbb{R}$; b) K este corpul $Q(\sqrt{2})$ avînd elemente de forma $a+b\sqrt{2}$ cu a, b numere raţionale.

9. Fie în $\mathbb{R}^3$ vectorii:

$$e_1=(1,0,0),\ e_2=(0,1,0),\ e_3=(0,0,1),\ e_4=(1,1,0),\ e_5=(1,0,1),\ e_6=(0,1,1)$$

a) Să se constate că (e_1,e_2,e_3), (e_2,e_3,e_4), (e_3,e_4,e_5), (e_4,e_5,e_6) sînt în baze în $\mathbb{R}^3$.

b) Ştiind că vectorul v are coordonatele (a,b,c) în baza (e_1,e_2,e_3) să i se determine coordonatele în celelalte trei baze.

10. Vectorului $u = (x,y,z)$ din $\mathbb{R}^3$ i se asociază vectorul $f(u) = (x+y, y-z)$ din $\mathbb{R}^2$.

a) Să se arate că f este operator liniar.

b) Să se precizeze matricea operatorului f în bazele canonice ale spaţiilor $\mathbb{R}^3$ şi $\mathbb{R}^2$.

c) Să se precizeze $Ker\,f$ şi $Im\,f$.

11. Vectorului $u = (x,y)$ din $\mathbb{R}^2$ i se asociază vectorul $f(u) = (x,x+y,y)$ din R^3. Să se dea răspunsurile la întrebările a), b), c) din exerciţiul precedent.

12. Vectorilor $x = (x_1,x_2)$ din $\mathbb{R}^2$ şi $y = (y_1,y_2,y_3)$ din $\mathbb{R}^3$ i se asociază numărul real

$$f(x,y) = x_1(y_1+2y_2-y_3) + x_2(2y_1+3y_2+y_3)$$

a) Să se arate că f este o formă biliniară.

b) Să se precizeze matricea asociată lui f în raport cu bazele canonice din $\mathbb{R}^2$ şi $\mathbb{R}^3$.

c) Alegînd în $\mathbb{R}^2$ baza canonică să se determine o bază a lui R^3 încît matricea asociată lui f să aibă formă diagonală.

13. Vectorului $x = (x_1,x_2,x_3)$ din $\mathbb{R}^3$ i se asociază

$p(x)=x_1^2+5x_2^2+12x_3^2-2x_1x_2+4x_1x_3-16x_2x_3$. Să se găsească forma canonică a formei pătratice p şi să se precizeze baza în care p admite acea formă canonică.

14. Să se determine polinomul P de grad cît mai mic încît $P(0)=3$, $P(1)=2$, $P(2)=3$, $P(3)=6$, $P(4)=11$.

BCU IAŞI/CENTRAL UNIVERSITY LIBRARY

SERII

§ 1. SERII NUMERICE

Fie şirul de numere reale:

$$(1) \quad a_1, a_2, a_3, \ldots, a_n, \ldots$$

Definiţia 3.1.

Se numeşte serie suma infinită:

$$(2) \quad a_1 + a_2 + a_3 + \ldots + a_n + \ldots$$

a_n se numeşte termenul general iar suma se mai prezintă şi în forma:

$$(3) \quad \sum_{n=1}^{\infty} a_n$$

Dată seria (2) sau (3) se formează şirul:

$$(4) \quad A_1 = a_1, \quad A_2 = a_1 + a_2, \ldots, A_n = a_1 + a_2 + \ldots + a_n, \ldots$$

numit şirul sumelor parţiale.

Definiţia 3.2.

Dacă $\lim_{n\to\infty} A_n = A$ (A finit sau nu) se spune că seria are limita A.

Dacă A este număr finit se spune că seria converge iar dacă A este

număr infinit se spune că seria diverge. Numărul A se numeşte suma seriei.

În legătură cu seriile se pun două probleme:

a) Să se cerceteze dacă seria este sau nu convergentă (să se stabilească natura seriei).

b) În caz de convergenţă să se calculeze suma.

Exemple:

1. O serie cunoscută deja este seria geometrică cu raţia $q \neq 1$.

$$a + aq + aq^2 + \ldots + aq^{n-1} + \ldots$$

Se ştie că $S_n = \dfrac{a - aq^n}{1 - q}$

Dacă $|q| < 1$ atunci $\lim\limits_{n\to\infty} S_n = \dfrac{a}{1-q}$ adică serie convergentă.

Dacă $|q| \geq 1$ seria diverge.

Pentru $a = 1$, $q = -1$ progresia dă $1 - 1 + 1 - 1 + \ldots$ cu sume parţiale 0 sau 1 deci caz în care suma nu există.

2. Seria cu $A_n = 1 + \dfrac{1}{1!} + \ldots + \dfrac{1}{n!}$ aşa cum s-a văzut în liceu este convergentă şi $1 + \sum\limits_{n=1}^{\infty} \dfrac{1}{n!} = e$

Proprietăţi:

1. Dacă într-o serie lăsăm la o parte un număr finit de termeni seria obţinută are aceeaşi natură ca şi seria iniţială.

În adevăr să lăsăm la o parte din (2) primii m termeni. Rămîne:

$$(5)\ a_{m+1} + a_{m+2} + \ldots + a_{m+k} + \ldots = \sum_{n=m+1}^{\infty} a_n$$

Să notăm cu $S_k = a_{m+1} + \ldots + a_{m+k}$ cu m fixat mai înainte. Evident că $A_{m+k} = A_m + S_k$ şi deci $S_k = A_{m+k} - A_m$. Trecînd la limită obţinem $S = A - A_m$ şi deci (5) converge şi are suma $A - A_m$.

Să presupunem acum că (5) converge şi $S_k \to S$, atunci avem

$S = A - A_m$ şi deci $A = S + A_m$.

Dacă notăm cu $\alpha_m = A - A_m$ constatăm imediat că pentru $m \to \infty$ $\alpha_m \to 0$.

Observaţii

1. Seria (5) se numeşte serie rest iar α_m-rest.

2. Putem enunţa proprietatea: dacă (2) converge atunci restul tinde la zero.

3. Dacă termenii seriei convergente (2) se înmulţesc cu acelaşi factor α se obţine o serie convergentă şi suma ei este αA.

4. Fie $\begin{matrix} A = a_1 + a_2 + \ldots + a_n + \ldots \\ B = b_1 + b_2 + \ldots + b_n + \ldots \end{matrix}$ şi $\alpha, \beta \in \mathbb{R}$ atunci seria $\alpha\left(\sum_{n=1}^{\infty} a_n\right) + \beta\left(\sum_{n=1}^{\infty} b_n\right)$ este tot convergentă. Notînd cu A_n, B_n, C_n sumele parţiale putem scrie:

$$C_n = (\alpha a_1 + \alpha b_1) + \ldots + (\alpha A_n + \beta b_n) = \alpha A_n + \beta B_n$$

şi deci $\lim_{n\to\infty} C_n = \alpha A + \beta B$.

5. Din relaţia $a_n = A_n - A_{n-1}$, rezultă $\lim_{n\to\infty} a_n = 0$. Condiţia nu este suficientă. În adevăr fie $\sum_{n=1}^{\infty} \frac{1}{\sqrt{n}}$. Avem evident $\frac{1}{\sqrt{n}} \to 0$ şi totuşi din faptul că

$$1 + \frac{1}{\sqrt{2}} + \frac{1}{3} + \ldots + \frac{1}{n} > n \cdot \frac{1}{\sqrt{n}} = \sqrt{n}$$

rezultă evident că seria diverge.

O altă serie cu o proprietate asemănătoare şi în acelaşi timp deosebit de importantă este *seria armonică*.

Aceasta este $\sum_{n=1}^{\infty} \frac{1}{n} = 1 + \frac{1}{2} + \frac{1}{3} + \ldots + \frac{1}{n} + \ldots$

Denumirea vine de la faptul că orice temen este media armonică a termenilor vecini.

Se mai observă deasemenea că:

$$\frac{1}{n}+\frac{1}{n+1}+\ldots+\frac{1}{2n-1}>n\cdot\frac{1}{2}n=\frac{1}{2}$$

şi deci grupînd termenii cîte 1,2,4,8... se obţine

$$1,\ \frac{1}{2}+\frac{1}{3};\ \frac{1}{4}+\frac{1}{5}+\frac{1}{6}+\frac{1}{7};\ldots.$$

Deoarece fiecare grupă este mai mare decît 1/2 rezultă că suma seriei este nămărginită şi deci rezultă că seria armonică este divergentă.

Profităm de prilej pentru a prezenta o serie cunoscută sub numele de serie armonică generalizată.

$$\sum_{n=1}^{\infty}\frac{1}{n^s}=1+\frac{1}{2^s}+\frac{1}{3^s}+\ldots+\frac{1}{n^s}+\ldots$$

unde $s\in\mathbb{R}$.

Dacă $s=1$ se obţine seria armonică prezentată mai înainte. Are loc afirmaţia: Pentru $s<1$ seria diverge iar pentru $s>1$ seria converge.

În adevăr dacă $s<1$ termenii seriei armonice generalizată sînt mai mari decît termenii corespunzători ai seriei armonice ei prin urmare şirul sumelor este nemărginit şi deci seria diverge.

Pentru $s>1$ să punem $s=1+t$, $t>0$.

Şi în cazul acestei serii se poate observa că:

$$\frac{1}{n^s}+\frac{1}{(n+1)^s}+\ldots+\frac{1}{(2n+1)^s}<n\cdot\frac{1}{n^s}=\frac{1}{n^t}$$

Făcînd grupări asemănătoare cu cele precedente se obţin grupe:

$$1;\ \frac{1}{2^s}+\frac{1}{3^s};\ \frac{1}{4^s}+\frac{1}{5^s}+\frac{1}{6^s}+\frac{1}{7^s};\ldots$$

Cum fiecare grupă , exceptînd-o pe prima, este mai mică decît $\frac{1}{2^t}$, $\frac{1}{4^t}=\frac{1}{2^{2t}}\ldots$ rezultă evident că suma seriei armonice generalizate este mai mică decît:

$$1+\frac{1}{2^t}+\frac{1}{2^{2t}}+\ldots=1+\frac{1}{1-\frac{1}{2^t}}$$

şi deci seria converge.

Definiţia 3.3.

Seria $(A)\ \sum_{n=1}^{\infty} a_n = a_1 + a_2 + \ldots + a_n + \ldots$ este serie cu termeni pozitivi dacă $a_n \geq 0$ pentru $n = 1, 2, \ldots$.

În cazul seriei cu termeni pozitivi şirul sumelor parţiale este monoton nedescrescător deoarece $A_{n+1} = A_n + a_{n+1} \geq A_n$.

Rezultă deci că seria (A) este convergentă şi suma va fi finită dacă şirul A_n este mărginit superior.

Problema precizării naturii unei serii cu termeni pozitivi poate fi soluţionată cu uşurinţă dacă vom indica procedee care să asigure mărginirea de care am vorbit.

În acest sens mai întîi vom prezenta cîteva teoreme numite teorema de convergenţă şi apoi unele condiţii suficiente (numite şi criterii) de convergenţă.

Teorema 3.1. (Prima teoremă de comparaţie).

Fie seriile cu termeni pozitivi $(A)\ \sum_{n=1}^{\infty} a_n$ şi $(B)\ \sum_{n=1}^{\infty} b_n$.

Dacă începînd cu un anumit termen fiecare termen al seriei (A) este mai mic decît termenul corespunzător al seriei (B) atunci din convergenţa seriei (B) rezultă convergenţa seriei (A) sau din divergenţa seriei (A) rezultă divergenţa seriei (B).

Demonstraţie: Putem presupune că $a_n \leq b_n$ pentru $n = 1, 2, \ldots$ deoarece lăsînd la o parte un număr finit de termeni natura seriei se conservă. Fie din nou A_n, B_n sumele parţiale corespunzătoare. Avem deci $A_n \leq B_n$.

Prin urmare dacă $B_n \to B$ rezultă că $A_n < B$ şi deci şirul sumelor

parţiale este mărginit adică seria (A) converge.

Se observă de asemeni că nemărginirea lui A_n implică nemărginirea lui B_n adică divergenţa seriei (A) implică şi divergenţa seriei (B).

Teorema 3.2. (Teorema a doua de comparaţie).

Dacă există $\lim_{n\to\infty} \frac{a_n}{b_n} = k \quad 0 < k < \infty$ atunci din convergenţa seriei (B) pentru $k < \infty$ urmează convergenţa seriei (A) iar din divergenţa seriei (A) pentru $k > 0$ rezultă divergenţa seriei (B).

Demonstraţie: Să presupunem că seria (B) converge şi fie k finit. Fie $\varepsilon > 0$. Atunci pentru n suficient de mare avem $\frac{a_n}{b_n} < k + \varepsilon$ şi deci $a_n < (k+\varepsilon) b_n$. Convergenţa seriei (B) implică convergenţa seriei $\sum (k+\varepsilon) b_n$. Urmează că din convergenţa seriei (B) rezultă mărginirea şirului A_n şi deci convergenţa seriei (A).

Partea a doua rezultă imediat prin reducere la absurd.

Teorema 3.3. (Teorema a treia de comparaţie).

Dacă de la un anumit indice N este îndeplinită inegalitatea $\frac{a_{n+1}}{a_n} \le \frac{b_{n+1}}{b_n}$ atunci convergenţa seriei (B) asigură convergenţa seriei (A) iar divergenţa seriei (A) divergenţa seriei (B).

Demonstraţie: Admitem, din nou, că relaţia din enunţ are loc pentru $n = 1, 2, \ldots$. Putem scrie:

$$\frac{a_2}{a_1} \le \frac{b_2}{b_1}, \quad \frac{a_3}{a_2} \le \frac{b_3}{b_2} \cdots \frac{a_n}{a_{n-1}} \le \frac{b_n}{b_{n-1}}$$

Prin înmulţirea membru cu membru se obţine $\frac{a_n}{a_1} \le \frac{b_n}{b_1}$ sau $a_n \le \frac{a_1}{b_1} b_n$ pentru $n = 1, 2, 3, \ldots$.

BCU IAŞI CENTRUL UNIVERSITAR LIBRARY

Convergenţa seriei (B) asigură convergenţa seriei $\sum_{n=1}^{\infty} \frac{a_1}{b_1} b_n$ şi prin urmare şi seria $\sum_{n=1}^{\infty} a_n$ este convergentă şi analog şi partea a doua.

Alegînd convenabil seriile (A) sau (B) se obţin condiţii suficiente de convergenţă numite criterii de convergenţă.

Reamintim că seria geometrică $\sum_{n=1}^{\infty} q^n$ este convergentă pentru $0 < q < 1$ iar pentru $q = 1$ divergentă.

Să folosim drept serie (B) seria geometrică. Vom obţine *criteriul lui Cauchy* sau *criteriul rădăcinii*. Fie $C_n = \sqrt[n]{a_n}$.

Teorema 3.4.

Dacă de la un n suficient de mare $C_n \le q$ unde $0 < q < 1$ atunci seria (A) converge iar dacă $C_n > 1$ seria (A) diverge.

Demonstraţie: Condiţia $C_n \le q \Rightarrow a_n \le q^n$ şi deci teorema 3.1 asigură convergenţa seriei (A).

Partea a doua asigură $a_n \ge 1$ şi deci comparînd cu $\sum 1$ rezultă divergenţa seriei (A).

Observaţie. Dacă $\lim_{n \to \infty} C_n = C$ (finită sau nu) se poate obţine forma limită a criteriului rădăcinii: Dacă $C < 1$ seria (A) converge, dacă $C > 1$ seria diverge iar dacă $C = 1$ nu putem decide (se spune că avem situaţia de dubiu).

Comparînd tot cu seria geometrică dar folosind teorema 3.3 se obţine *criteriul lui D'Alembert* sau *criteriul raportului*.

Teorema 3.5.

Fie $d_n = \frac{a_{n+1}}{a_n}$. Dacă pentru n suficient de mare $d_n \le q \quad 0 < q < 1$

atunci (A) converge iar dacă $d_n \geq 1$ urmează că seria (A) diverge.

Demonstraţie: În adevăr dacă $b_n = q^n$ din teorema 3.3 rezultă imediat concluzia:

Observaţie: În forma limită: Fie $\lim_{n\to\infty} d_n = d$

Dacă $d < 1$ seria (A) converge, dacă $d > 1$ seria (A) diverge iar pentru $d = 1$ avem caz de dubiu.

Un criteriu cu putere mai mare în privinţa înlăturării dubiului este *criteriul lui Raabe.*

Teorema 3.6.

Fie $r_n = n\left(1 - \frac{a_{n+1}}{a_n}\right)$.

Dacă pentru n suficient de mare $r_n \geq n$, $r > 1$ atunci seria (A) converge iar dacă $r_n \leq 1$ seria (A) diverge.

Demonstraţie: Presupunem că pentru n suficient de mare avem: $n\left(1 - \frac{a_{n+1}}{a_n}\right) \geq r > 1$ sau $\frac{a_{n+1}}{a_n} \leq 1 - \frac{r}{n}$.

Fie t, $1 < t < r$ atunci avem:

$$(*) \quad \lim_{n\to\infty} \frac{\left(1 - \frac{1}{n}\right)^t - 1}{-\frac{1}{n}} = t$$

Într-adevăr să pornim de la limita cunoscută:

$$\lim_{\alpha\to 0} \frac{\ln(1+\alpha)}{\alpha} = 1$$

şi să notăm $\left(1 - \frac{1}{n}\right)^t - 1 = \beta$. Evident că pentru $n \to \infty$ $\beta \to 0$. Avem

$\left(1-\frac{1}{n}\right)^t = \beta + 1$ şi logaritmînd

$$t \ln\left(1-\frac{1}{n}\right) = \ln(1+\beta)$$

Deci avem

$$\frac{\left(1-\frac{1}{n}\right)^t - 1}{-\frac{1}{n}} = \frac{\beta}{-\frac{1}{n}} = \frac{\beta}{\ln(1+\beta)} \cdot \frac{\ln\left(1-\frac{1}{n}\right)^t}{-\frac{1}{n}}$$

şi deci are loc (*).

Revenind acum la demonstraţie, pentru n suficient de mare avem:

$\frac{\left(1-\frac{1}{n}\right)^t - 1}{-\frac{1}{n}} < r$ sau $\left(1-\frac{1}{n}\right)^t > 1 - \frac{r}{n}$ deci şi:

$$\frac{a_{n+1}}{a_n} < \left(1-\frac{1}{n}\right)^t$$

Aceasta se scrie:

$$\frac{a_{n+1}}{a_n} < \frac{(n-1)^t}{n^t} = \frac{\frac{1}{n^t}}{\frac{1}{(n-1)^t}}$$

În felul acesta se pun în evidenţă termeni ai seriei armonice generalizate cu $t > 1$.

Cum aceasta este convergentă pentru $t > 1$, avînd în vedere teorema 3.3 rezultă că seria (A) converge.

În cazul $n\left(1-\frac{a_{n+1}}{a_n}\right) \le 1$ rezultă $\frac{a_{n+1}}{a_n} \ge \frac{n-1}{n} = \frac{\frac{1}{n}}{\frac{1}{n-1}}$ şi comparaţia se face cu seria armonică şi deci tot din teorema 3.3 obţinem că seria (A) diverge.

BCU IAŞI CENTRAL UNIVERSITY LIBRARY

În felul acesta se pun în evidenţă termeni ai seriei armonice generalizate cu $t>1$.

Cum aceasta este convergentă pentru $t>1$, avînd în vedere teorema 3.3 rezultă că seria (A) converge.

În cazul $n\left(1-\frac{a_{n+1}}{a_n}\right)\le 1$ rezultă $\frac{a_{n+1}}{a_n}\ge\frac{n-1}{n}=\frac{\frac{1}{n}}{\frac{1}{n-1}}$ şi comparaţia se face cu seria armonică şi deci tot din teorema 3.3 obţinem că seria (A) diverge.

Observaţie: Pentru a obţine forma limită fie $\lim\limits_{n\to\infty} r_n = r$.

Dacă $r>1$ seria (A) converge, dacă $r<1$ seria diverge iar dacă $r=1$ - dubiu.

Exemple:

Să se precizeze natura seriilor:

1. $\sum\limits_{n=1}^{\infty}\left(1-\frac{1}{n}\right)^{n^2}$

Aplicăm criteriul lui Cauchy $c_n=\left(1-\frac{1}{n}\right)^n$, $\lim\limits_{n\to\infty} c_n=\frac{1}{e}<1$ deci seria este convergentă.

2. $1+\sum\limits_{n=1}^{\infty}\frac{x^n}{n!}\quad (x>0)$

Folosind criteriul lui D'Alembert $d_n=\frac{a_{n+1}}{a_n}=\frac{x}{n+1}$ $\lim\limits_{n\to\infty} d_n=0$ şi deci $(\forall)x$ seria converge.

3. $\sum\limits_{n=1}^{\infty} n!\left(\frac{x}{n}\right)^n$ $(x>0)$. Acum $d_n=\frac{x}{\left(1+\frac{1}{n}\right)^n}$ şi $d=\frac{x}{e}$.

BCU IAŞI CENTRAL UNIVERSITY LIBRARY

Prin urmare:

1. dacă $x < e \Rightarrow d < 1$ şi seria converge;
2. dacă $x > e \Rightarrow d > 1$ seria diverge şi
3. dacă $x = e$ caz de dubiu.

4. $1 + \sum_{n=1}^{\infty} \frac{(2n-1)!!}{2n!!} \cdot \frac{1}{2n+1}$

Aplicînd criteriul lui D'Alembert $d_n = \frac{(2n-1)^2}{2n(2n+1)}$ rezultă $d = 1$ şi deci dubiu. Să aplicăm şi criteriul lui Raabe.

$$r_n = n\left(1 - \frac{(2n-1)^2}{2n(2n+1)}\right) = \frac{6n-1}{2(2n+1)}, \quad r_n \to \frac{3}{2} > 1$$

şi deci seria este convergentă.

Nu toate seriile sînt cu termeni pozitivi. Seriile cu termeni pozitivi şi negativi se numesc serii cu termeni oarecare iar seria $a_1 - a_2 + a_3 - a_4 + \ldots + (-1)^{n-1} a_n + \ldots$ (cu toţi $a_i > 0$) se numeşte serie alternantă.

Pentru astfel de serii consideraţiile de mai sus nu mai sînt valabile (şirul seriilor parţiale nu mai este, în general, monoton crescător).

Pentru serii cu termeni oarecare are loc criteriul general de convergenţă a lui Cauchy.

Teorema 3.7.

Condiţia necesară şi suficientă ca seria (A) să conveargă este ca pentru orice $\varepsilon > 0$ să existe N aşa încît pentru $n > N$

$$|A_{n+m} - A_n| = |a_{n+1} + \ldots + a_{n+m}| < \varepsilon$$

să aibă loc pentru orice $m = 1, 2, 3, \ldots$

rezultă de aici că $|a_{n+1}| < \varepsilon$ adică $a_{n+1} \to 0$ cînd $n \to \infty$.

Acest criteriu general este nepractic.

BCU IASI CENTRAL UNIVERSITY LIBRARY

Un pas în rezolvarea problemei determinării naturii unei serii oarecare îl face teorema următoare:

Teorema 3.8.

Dacă $\sum_{n=1}^{\infty} |a_n|$ este convergentă atunci $\sum_{n=1}^{\infty} a_n$ este convergentă.

Demonstraţie: Să folosim criteriul general de convergenţă a lui Cauchy. Fie $\varepsilon > 0$ şi $\sum_{n=1}^{\infty} |a|$ convergentă; atunci pentru $n > N$ şi $m = 1, 2, 3, \ldots$ avem $||a_{n+1}| + \ldots + |a_{n+m}|| < \varepsilon$ sau

$$|a_{n+1} + \ldots + a_{n+m}| < |a_{n+1}| + \ldots + |a_{n+m}| < \varepsilon$$

pentru $n > N$ şi $m = 1, 2, \ldots$ şi deci $\sum_{n=1}^{\infty} a_n$ este convergentă.

Definiţia 3.4.

Seria $\sum_{n=1}^{\infty} a_n$ pentru care seria $\sum_{n=1}^{\infty} |a_n|$ este convergentă se numeşte *absolut convergentă*.

Pentru seriile cu termeni oarecare în cazul în care $\sum_{n=1}^{\infty} a_n$ este convergentă iar seria modulelor termenilor este divergentă se spune că seria $\sum_{n=1}^{\infty} a_n$ este serie semiconvergentă.

Iată un exemplu:

$$\frac{1}{1} - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \ldots\ldots$$

este o serie convergentă în timp ce seria modulelor termenilor este o serie divergentă.

Pentru seriile alternante are loc afirmaţia:

BCU IAŞI CENTRAL UNIVERSITY LIBRARY

Teorema 3.9.

Dacă $\lim_{n\to\infty} a_n = 0$ şi $|a_1| > |a_2| \ldots\ldots$ atunci seria este convergentă. (Fără demonstraţie).

§ 2. SERII DE FUNCŢII

Definiţia 3.5.

O serie de forma:

(1). $\sum_{n=0}^{\infty} a_n x^n = a_0 + a_1 x + \ldots + a_n x^n + \ldots$, $a_n, x \in \mathbb{R}$ se numeşte serie de puteri (serie întreagă). Numerele a_n se numesc coeficienţii seriei. Pentru anumite valori particulare ale lui x se obţin serii numerice care pot fi convergente sau nu.

Mulţimea valorilor lui x în care seria considerată este convergentă se numeşte *domeniul de convergenţă* al seriei. Acest domeniu poate fi finit sau nu. Pentru determinarea acestui domeniu esenţială este:

Teorema 3.10.

Dacă există un număr A şi un x_0 aşa fel ca $|a_n x_0^n| < A$ pentru orice n, atunci seria $\sum_{n=0}^{\infty} a_n x^n$ este absolut convergentă pentru toate valorile lui x care verifică relaţia:

$$|x| < |x_0|$$

Demonstraţie: Din convergenţa seriei (1) pentru x_0 urmează că orice sumă parţială este mărginită şi deci şi $|a_n x_0^n| < A$ pentru un A oarecare. Fie x aşa fel ca $|x| < |x_0|$.

Putem scrie: $\sum_{n=0}^{\infty} |a_n x^n| = |a_0| + |a_1 x| + \ldots + |a_n x^n| + \ldots$ iar

$|a_n x^n| = |a_n x_0^n| \left|\frac{x}{x_0}\right|^n < A\left|\frac{x}{x_0}\right|^n$ şi termenii sînt mai mici ca termenii corespunzători ai progresiei geometrice cu raţia $\left|\frac{x}{x_0}\right| < 1$. Convergenţa seriei geometrice asigură seriei (1) absoluta convergenţă în $x < x_0$.

Condiţia $|x| < |x_0|$ echivalentă cu $-|x_0| < x < |x_0|$ arată că în general domeniul de convergenţă este un interval cu centrul în origine şi deci are forma *(-R,R)*. Pentru $x = \pm R$ seria poate converge sau nu iar pentru $|x| > R$ seria este divergentă. Numărul R astfel definit se numeşte *raza de convergenţă* a seriei (1).

Observaţie: Termenul „rază" se justifică suplimentar în cazul în care x şi eventual a_n sînt numere complexe. Consideraţii analoge asigură convergenţa pentru $|x| < R$ şi divergenţa pentru $|x| > R$, cazurile de dubiu şi cele de „semi-convergenţă" survenind cînd numărul complex x se plasează pe un cerc de rază R.

Pentru calculul razei de convergenţă observăm că aplicînd criteriile lui D'Alembert sau Cauchy se obţine:

$$R = \frac{1}{\lim\limits_{n\to\infty} \sqrt[n]{|a_n|}} \quad \text{sau} \quad R = \lim_{n\to\infty} \left|\frac{a_n}{a_n + 1}\right|$$

Exemple

Seria $\sum\limits_{n=0}^{\infty} x^n$ are raza de convergenţă $R = 1$ şi deci intervalul de convergenţă (-1,1), iar seria $\sum\limits_{n=0}^{\infty} \frac{x^n}{n!}$ are raza infinită şi deci este convergentă pe toată axa reală. Aflăm de asemenea că $\sum\limits_{n=0}^{\infty} n!\, x^n$ are raza de convergenţă egală cu zero.

Dată o serie de puteri se pune în evidenţă posibilitatea definirii unei

funcţii definită pe domeniul de convergenţă şi cu valori în $\mathbb{R}$. Se obişnuieşte să se spună că această funcţie este suma seriei (1). Pentru a putea prezenta proprietăţi ale acestei funcţii vom pune problema într-un cadru mult mai general.

În acest scop vom considera şirul de funcţii reale $f_1, f_2, \ldots, f_n, \ldots$ toate definite pe domeniul D cu valori în $\mathbb{R}$ şi vom admite că pentru orice $x \in D$ şirul:

(2). $f_1(x), f_2(x), \ldots, f_n(x), \ldots$ are limită. Obţinem astfel o funcţie $f: D \to R$ pe care o vom numi funcţia limită şi vom scrie:

(3). $f(x) = \lim_{n \to \infty} f_n(x)$

Mai mult deci, dacă avem şirul de funcţii $u_1, u_2, \ldots, u_n, \ldots$ sîntem conduşi la considerarea seriei:

(4). $\sum_{n=1}^{\infty} u_n(x) = u_1(x) + u_2(x) + \ldots + u_n(x) + \ldots$ cu sumele parţiale date de $f_n(x) = u_1(x) + u_2(x) + \ldots + u_n(x)$.

Dacă admitem că seria (4) converge în orice x din domeniul D de definiţie al funcţiei $u_1, \ldots, u_n$ se obţine de fapt funcţia din (3).

De obicei interesează dacă anumite proprietăţi ale funcţiei $f_1, f_2, \ldots$ se transmit funcţiei f.

Definiţia 3.6.

Dacă:

1. Şirul (2) are în D limita $f(x)$.

2. Pentru fiecare $\varepsilon > 0$ $(\exists)$ $N(\varepsilon)$ astfel încît pentru $n > N(\varepsilon)$, $|f_n(x) - f(x)| < \varepsilon$ să aibă loc pentru orice $x \in D$. Vom spune că (2) *converge uniform* la $f(x)$.

Exemplu:

Orice serie de puteri este uniform convergentă pe intervalul său de convergenţă.

În adevăr: să notăm cu $r_n(x) = a_{n+1}x^{n+1} + a_{n+2}x^{n+2} + \ldots$ expresie numită *restul seriei*.

Fie $\varepsilon > 0$ şi x un punct oarecare din intervalul de convergenţă $(-R, R)$. Să arătăm că putem determina un rang $N(\varepsilon)$ independent de x, aşa fel ca pentru $x > N(\varepsilon)$ restul seriei să îndeplinească condiţia

$$|r_n(x)| = |a_{n+1}x^{n+1} + \ldots| < \varepsilon$$

Să alegem un x_0 din $(-R, R)$ aşa fel ca

$$|x| < |x_0| < R$$

Din convergenţa seriei $\sum_{n=0}^{\infty} a_n x_0^n$ în $(-R, R)$ restul $r_n(x_0)$ îndeplineşte condiţia:

$$|r_n(x_0)| = |a_{n+1}x_0^{n+1} + \ldots| < \varepsilon$$

începînd cu un anumit rang $n > N(\varepsilon)$.

Dar $|a_n x^n| < |a_n x_0^n|$ aşa că putem scrie $|r_n(x)| < \varepsilon$ de la un rang $n > N(\varepsilon)$ oricare ar fi $x \in (-R, R)$.

Teorema 3.11.

Dacă $u_1, u_2, u_3, \ldots, u_n, \ldots$ sînt definite şi continui pe $[a, b]$ şi seria (2) converge uniform la f atunci şi f este continuă pe $[a, b]$.

Demonstraţie: Fie $x_0 \in [a, b]$ atunci din $f(x) = f_n(x) + r_n(x)$ rezultă că $f(x_0) = f_n(x_0) + r_n(x_0)$. De aici decurge:

$$|f(x) - f(x)| \le |f_n(x) - f_n(x)| + |r_n(x)| + |r_n(x_0)|$$

Fie $\varepsilon > 0$, se poate fixa n aşa încît $|r_n(x)| < \frac{\varepsilon}{3}$ $(\forall) x \in D$.

Atunci din continuitatea funcţiei f_n rezultă că dacă $|x - x_0| < \delta$ decurge $|f_n(x) - f_n(x)| < \frac{\varepsilon}{3}$ şi deci în concluzie $|f(x) - f(x_0)| < \varepsilon$, adică tocmai teorema.

Consecinţă: Dacă f este suma seriei $\sum_{n=0}^{\infty} a_n x^n$ unde $x \in (-\mathbb{R}, \mathbb{R})$, f este o funcţie continuă.

Au deasemenea loc următoarele afirmaţii:

Teorema 3.12.

a) Dacă $\lim_{x \to x_0} u_n = c_n$, $n = 1,2,..$ şi (2) converge uniform la f, atunci $\lim_{n \to \infty} c_n = c$ şi $\lim_{x \to x_0} f(x) = c$.

b) Dacă u_n, $n = 1,2,3,...$ este continuă pe $[a,b]$ şi (2) converge uniform la f atunci are loc:

$$\int_a^b f(x)\,dx = \int_a^b u_1(x)\,dx + \int_a^b u_2(x)\,dx + \ldots$$

c) Dacă u_n, $n = 1,2,...$ sînt definite pe $[a,b]$ şi există derivate $u_n'(x)$ şi (2) converge uniform la f atunci

$$\sum_{n=1}^{\infty} u_n'(x) = f'(x)$$

Consecinţă: O serie de puteri se poate deriva termen cu termen şi seria obţinută este convergentă şi aceeaşi rază de convergenţă.

Observaţie: Fie seria de puteri:

$$f(x) = a_0 + a_1 x + a_2 x^2 + \ldots + a_n x^n + \ldots$$

Conform consecinţei precedente putem deriva termen cu termen şi vom obţine serii noi cu aceeaşi rază de convergenţă.

Procedînd astfel obţinem:

$$f'(x) = a_1 + 2a_2 x + \ldots$$
$$f''(x) = 2a_2 + \ldots\ldots$$
$$\ldots\ldots\ldots\ldots$$

De unde făcînd $x = 0$:

$$f(0)=a_0,\ f'(x)=a_1,\ f''(x)=2a_2,\ldots$$

şi deci f se scrie:

$$(5)\quad f(x)=f(0)+\frac{x}{1!}f'(0)+\frac{x^2}{2!}f''(0)+\ldots+\frac{x^n}{n!}f^{(n)}(0)+\ldots$$

numită seria lui Mac-Laurin. Deci o funcţie f indefinit derivabilă se dezvoltă în serie de puteri după formula lui Mac-Laurin.

Definiţia 3.7.

O funcţie care admite o dezvoltare în serie de puteri se numeşte *funcţie analitică.*

În practică alături de seria Mac-Laurin este folosită frecvent o serie care o generalizează numită seria Taylor.

Vom prezenta mai întîi formula lui Taylor pentru un polinom de gradul n (mai precis funcţia polinomială ataşată).

Fie $f(x)=a_0+a_1x+a_2x^2+\ldots+a_nx^n$. Dorim să găsim coeficienţii $A_0,A_1,\ldots,A_n$ aşa ca:

$$f(x)=A_0+A_1(x-x_0)+A_2(x-x_0)^2+\ldots+A_n(x-x_0)^n$$

unde x_0 este o valoare reală fixată.

Evident $f(x)=A_0$, $f'(x)=A_1$, $f''(x_0)=2A_2,\ldots,f^{(n)}(x_0)=n!A_n$.

Prin urmare avem:

$$f(x)=f(x)+\frac{f'(x)}{1!}(x-x_0)+\frac{f''(x_0)}{2!}(x-x_0)^2+\ldots+\frac{f^{(n)}(x_0)}{n!}(x-x_0)^n$$

numită formula lui Taylor pentru polinomul f.

Dacă f notează o funcţie indefinită derivabilă şi x_0 aparţine domeniului ei de definiţie are loc formula.

$$(6)\quad f(x)=f(x_0)+\frac{f'(x_0)}{1!}(x-x_0)+\ldots+\frac{f''(x_0)}{n!}(x-x_0)^n+r_n(x)$$

numită formula lui Taylor pentru funcţia f în jurul lui x_0.

BCU CLUJ CENTRAL UNIVERSITY LIBRARY

Observaţii:

1. Dacă $x_0 = 0$ din (6) se găseşte (5).

2. Restul $r_n(x)$ are diverse expresii şi are importanţă esenţială în calculul valorii funcţiei f în punctul x.

O expresie des întîlnită pentru rest este formula lui Lagrange:

$$r_n(x) = \frac{f^{(n+1)}(\theta x)}{(n+1)!} x^{n+1} \quad \theta \in (0,1)$$

3. Dacă în (6) ne oprim la n se obţine o serie de forma $\sum_{n=0}^{\infty} a_n (x - x_0)^n$, $a_n = \frac{f^{(n)}(x_0)}{n!}$ numită serie Taylor.

Fiind tot o serie de puteri înseamnă că şi aici se pune problema determinării razei de convergenţă.

Raţionînd analog ca mai sus asupra serici:

$$|C_0| + |C_1|\,|x-a| + |C_2|\,|x-a|^2 + \ldots + |C_n|\,|x-a|^n + \ldots..$$

se obţine R şi deci convergenţa va fi asigurată pentru $|x-a| < R$ adică $a - R < x < a + R$ şi deci în acest caz domeniul de convergenţă este de forma: $(a-R, a+R)$. De exemplu: Fie $\sum_{n=1}^{\infty} \frac{(x+1)^n}{n2^n}$. Obţinem $\lim_{n\to\infty} \frac{|u_{n+1}|}{|u_n|} = \frac{|x+1|}{2}$ şi deci seria converge pentru $\frac{|x+1|}{2} < 1$, de unde $-3 < x < 1$.

Exemple de funcţii dezvoltate în serii de puteri.

1) $e^x = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \ldots, + \frac{x^n}{n!} + \ldots \quad x \in \mathbb{R}$

2) $\frac{1}{1-x} = 1 + x + x^2 + \ldots, + x^n + \ldots \quad |x| < 1$

3) $\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \ldots, (-1)^{n-1}\frac{x^n}{n} + \ldots \quad |x| < 1$

4) $\sin x = x - \frac{1}{3!}x^3 + \frac{1}{5!}x^5 + \ldots + (-1)^{2n+1}\frac{1}{(2n+1)}x^{2n+1} + \ldots \quad x \in R$

5) $\cos x = 1 - \frac{1}{2!}x^2 + \frac{1}{4!}x^4 + \ldots + (-1)^{2n}\frac{1}{(2n)!}x^{2n} + \ldots \quad x \in R$

BCU IASI CENTRUL UNIVERSITAR LIBRARY

Deoarece seria din exemplul 1) converge pentru $x \in R$ trecînd x în $-x$ se obţine tot serie convergentă pe R. Avem:

6) $e^{-x} = 1 - \frac{x}{1!} + \frac{x^2}{2!} - \frac{x^3}{3!} + \frac{x^4}{4!} + \ldots + (-1)^n \frac{x^n}{n!} + \ldots$

Scăzînd şi adunînd membru cu membru egalităţile din exemplele 1) şi 6) obţinem două noi serii convergente pe $\mathbb{R}$.

7) $\frac{1}{2}(e^x - e^{-x}) = x + \frac{1}{3!}x^3 + \frac{1}{5!}x^5 + \ldots + \frac{1}{(2n+1)!}x^{2n+1} + \ldots$

8) $\frac{1}{2}(e^x + e^{-x}) = 1 + \frac{1}{2!}x^2 + \frac{1}{4!}x^4 + \ldots + \frac{1}{(2n)!}x^{2n} + \ldots$

Datorită proprietăţilor lor aceste funcţii limită din 7) şi 8) au fost numite sinus hiperbolic respectiv cosinus hiperbolic şi notate cu sh şi respectiv cu ch. Deci:

$$shx = \frac{1}{2}(e^x - e^{-x})$$

$$chx = \frac{1}{2}(e^x + e^{-x})$$

De altfel aceste funcţii au fost introduse în manualu de Analiză din clasa a-XI a cînd s-au evidenţiat şi unele proprietăţi ale lor.

Deşi cele ce urmează nu sînt justificate vă propunem un mic rabat de la rigoare şi să admitem formal extinderea definiţiei lui e^t de la valori reale $t = x$ la valori pur imaginare $t = \pm ix$. Înlocuind x prin ix şi apoi prin $-ix$ în egalitatea de la exemplul 1) obţinem:

$$e^{ix} = \cos x + i \sin x$$
$$e^{-ix} = \cos x - i \sin x$$

iar de aici

$$\sin x = \frac{1}{2i}(e^{ix} - e^{-ix})$$

$$\cos x = \frac{1}{2}(e^{ix} + e^{-ix})$$

Aceste două formule sînt cunoscute sub denumirea de formulele lui Euler.

BCU IASI/CLUJ UNIVERSITY LIBRARY

Întrebări de autocontrol şi exerciţii

1. Ce este o serie?
2. Ce este o serie convergentă?
3. De ce studiem cu atenţie aparte seriile cu termeni pozitivi?
4. Ce sînt criteriile de convergenţă?
5. Cum abordăm convergenţa seriilor cu termeni oarecare? Ce dificultăţi apar?
6. Ce este raza de convergenţă? Ce este domeniul de convergenţă?
7. Cum se motivează necesitatea introducerii uniformei convergenţe.
8. Ce problemă rezolvă formulele Taylor şi Mac-Laurian?
9. Precizaţi natura seriilor:

a) $\sum_{n=1}^{\infty} \frac{2n-1}{(\sqrt{2})^n}$, b) $\sum_{n=1}^{\infty} \left(\frac{n+1}{2n-1}\right)^n$,

c) $\sum_{n=1}^{\infty} \frac{1}{(3n-2)(3n+1)}$, d) $\sum_{n=1}^{\infty} \frac{a^n n!}{n^n}$ $(a>0)$

10. Să se dezvolte $f(x) = \frac{1}{x^2+3x+2}$ după puterile lui $x+4$.

11. Dezvoltaţi în serie Mac-Laurin $f(x) = \frac{3}{(1-x)(1+2x)}$.

12. Stabiliţi domeniul de convergenţă pentru seriile

a) $\sum_{n=1}^{\infty} \frac{1}{n!x^n}$;

b) $\sum_{n=1}^{\infty} \frac{n}{n+1}x^n$;

c) $\sum_{n=1}^{\infty} \frac{(x-2)^2}{(2n-1)2^n}$.

BCU IAŞI CENTRUL UNIVERSITAR LIBRARY

CAPITOLUL IV

FUNCŢII DE MAI MULTE VARIABILE

§ 1. GENERALITĂŢI

Foarte multe probleme practice şi mai ales din domeniul cercetării economice şi sociale impun folosirea unor funcţii care depind de mai multe variabile. Funcţiile de mai multe variabile se definesc analog cu funcţiile de o variabilă. Fie $D \in R^n$. Orice aplicaţie $f: D \to R$ se numeşte funcţie de n variabile. Vom nota ca de obicei imaginea punctului $(x_1, x_2, ..., x_n)$ prin $f(x_1, x_2, ..., x_n)$, iar mulţimea D o vom numi domeniul de definiţie.

Observaţie: Pentru înţelegerea celor ce urmează precizăm că domeniul D din R^n poate fi dat în foarte multe moduri în funcţie de n şi de problemă. De cele mai multe ori mulţimile de puncte se dau prin punctele unui paralelipiped dreptunghic dat prin condiţii de forma:

$$a_1 \le x_1 \le b_1 \ , \ a_2 \le x_2 \le b_2 \ , \ \dots \ , \ a_n \le x_n \le b_n$$

Pentru $n = 2$ se obţine dreptunghiul; $n = 3$ paralelipipedul dreptunghic ş.a.m.d

Dacă se exclud egalităţile se obţin domeniile deschise, analoagele intervalelor deschise.

Pentru $n = 2$ domeniul deschis se poate da prin $a < x < b$, $c < y < d$.

Numind vecinătatea unui punct $M_0(x_1^0, x_2^0, ..., x_n^0)$ orice paralelipiped deschis rezultă că o vecinătate simetrică sau cub se poate descrie prin: $x_1^0 - \delta < x_1 < x_1^0 + \delta$, ... , $x_n^0 - \delta < x_n < x_n^0 + \delta$ ($\delta > 0$). Un punct M cu coordonatele satisfăcînd condiţiile precedente se numeşte punct interior vecinătăţii.

Domeniile şi vecinătăţile pot fi descrise şi cu ajutorul sferei n—dimensionale.

Dacă punctul $(x_1, x_2, ..., x_n)$ îl notăm cu M atunci $z = f(x_1, .., x_n)$ se notează cu $z = f(M)$ numindu—se chiar funcţie de punct.

Definiţia 4.1.

Se spune că funcţia $f(x_1, ..., x_n) = f(M)$ are limita A pentru cazul în care M tinde la $M_0(a_1, ..., a_n)$ dacă oricum am alege $M_1, M_2, ..., M_p, ...$ ($M_i \neq M_0$) un şir de puncte care să conveargă la M_0 şirul $f(M_1)$, $f(M_2)$, ... , $f(M_p)$, ... converge la A.

Acest fapt se notează prin:

$$A = \lim_{M \to M_0} f(M)$$

sau

$$A = \lim_{\substack{x_1 \to a_1 \\ \cdots\cdots \\ x_n \to a_n}} f(x_1, ..., x_n).$$

Deci pentru cazul a două variabile dată funcţia $z = f(x,y)$ spunem că $f(x,y)$ tinde către o limită A, dacă pentru orice număr pozitiv ϵ, oricît de mic, putem determina un număr δ, astfel încît pentru $|x-a| < \delta$, $|y-b| < \delta$ să rezulte:

$$|f(x,y) - A| < \epsilon$$

Vom nota ca şi mai sus prin $\lim\limits_{\substack{x \to a \\ y \to b}} f(x,y) = A$.

Definiţia 4.2.

Se spune că funcţia $z = f(M)$ este continuă în punctul M_0 din D dacă

$\lim\limits_{M \to M_0} f(M) = f(M_0)$.

Observaţie: Atunci cînd scriem $\lim\limits_{\substack{x \to a \\ y \to b}} f(x,y) = f(a,b)$ vom considera că x şi y tind independent unul de altul la a, b. Acest fapt este valabil în general.

Această observaţie se impune deoarece dacă îl fixăm pe y şi facem pe $x \to a$ obţinem $\lim\limits_{x \to a} f(x,y) = \varphi(y)$. Făcînd apoi $y \to b$ se obţine $\lim\limits_{y \to b} (\varphi(y)) = \lim\limits_{y \to b} \lim\limits_{x \to a} f(x,y)$. Analog putem obţine $\lim\limits_{x \to a} \lim\limits_{y \to b} f(x,y)$. Asemenea limite se numesc limite iterate (repetate) pe cînd cealaltă se numeşte limită dublă (globală).

În general limitele iterate nu sînt egale şi chiar dacă ar fi limita globală poate nici să nu existe.

Are loc afirmaţia: dacă există limita globală $A = \lim\limits_{\substack{x \to a \\ y \to b}} f(x,y)$ şi pentru orice y există limita în raport cu x, $\varphi(y) = \lim\limits_{x \to a} f(x,y)$ atunci există limita iterată: $\lim\limits_{y \to b} \varphi(y) = \lim\limits_{y \to b} \lim\limits_{x \to a} f(x,y)$ şi este egală cu limita globală.

Teorema 4.1.

Dacă $z = f(x,y)$ este continuă atunci ea este continuă în raport cu fiecare variabilă separat.

Demonstraţie: Considerăm cazul $n = 2$. Dacă funcţia este continuă în (a,b) atunci făcînd $y = b$ avem:

$\lim\limits_{x \to a} f(x,b) = f(a,b)$ şi făcînd $x = a$

$\lim\limits_{y \to b} f(a,y) = f(a,b)$.

Reciproca nu este adevărată. O funcţie continuă separat în raport cu x şi separat în raport cu y nu este continuă ca funcţie de două variabile.

Aşa de exemplu: $f(x,y) = \dfrac{x-y}{x+y}$ este continuă ca funcţie de x în

origine şi $\lim_{x\to 0} f(x,y) = -1$. Ca funcţie de y de asemenea şi $\lim_{y\to 0} f(x,y) = 1$ pe cînd în $(0,0)$ $f(x,y)$ nu există.

§ 2. DERIVATE PARŢIALE. EXTREME

Proprietăţile învăţate la funcţiile de o variabilă se extind la cazul a n—variabile considerînd formulări adecvate. Această analogie ne sugerează să punem şi problema derivării.

Fie $z = f(x,y)$ definită pe (D) $\begin{cases} a_1 \le x \le b_1 \\ a_2 \le y \le b_2 \end{cases}$ şi $(a,b) \in D$.

Dacă privim y ca o constantă sau facem $y = b$, z devine o funcţie de o singură variabilă. Derivata acesteia dacă există în $x = a$ se numeşte derivata parţială în raport cu x a funcţiei $f(x,y)$ în (a,b).

Se notează cu $\frac{\partial f}{\partial x}$, dacă $x \in [a_1,b_1]$ şi $\frac{\partial f}{\partial x}$ există pentru orice x se obţine o nouă funcţie, funcţia derivată parţială de ordinul unu în raport cu x. Analog se poate proceda şi cu y.

Exemple: Să se calculeze derivatele parţiale pentru:

1) $f(x,y) = 3axy - x^3 - y^3$,

$$\frac{\partial f}{\partial x} = 3ay - 3x^2$$

$$\frac{\partial f}{\partial y} = 3ax - 3y^2$$

2) $f(x,y) = \frac{x^2 y}{x^2+y^2}$,

$$\frac{\partial f}{\partial x} = \frac{2xy^3}{(x^2+y^2)^2}$$

$$\frac{\partial f}{\partial y} = \frac{x^2(x^2-y^2)}{(x^2+y^2)^2}$$

Pentru $n = 3, 4, \ldots$ raţionîndu-se la fel se obţin derivate parţiale în raport cu fiecare variabilă.

Derivatele parţiale fiind şi ele funcţii de mai multe variabile se pot defini derivatele lor parţiale (dacă există). Se obţin astfel derivate parţiale de ordinul doi ale funcţiei iniţiale iar derivarea lor ne va conduce la derivate parţiale de ordinul trei ş.a.m.d. Derivatele parţiale de ordin superior se obţin pe o cale inductivă asemănător cu situaţia de la funcţiile de o singură variabilă.

Se obţin deci $\frac{\partial^2 f}{\partial x^2}$, $\frac{\partial^2 f}{\partial y^2}$, $\frac{\partial^2 f}{\partial x \partial y}$, $\frac{\partial^2 f}{\partial x \partial y}$. Ultimele două se numesc şi derivate parţiale mixte. În general derivatele parţiale mixte nu coincid, egalitatea lor are loc numai dacă acestea sînt continue ca funcţii de x şi y.

Fie $f: (a, b) \to \mathbb{R}$ o funcţie derivabilă în $x_0 \in (a,b)$.

După cum ştim urmărind să evaluăm creşterea unei funcţii, f, de o variabilă am fost conduşi la:

$$f(x) - f(x_0) \approx (x - x_0) f'(x_0)$$

Adică în vecinătatea lui x_0 $\Delta f \approx f'(x_0)\Delta x$ unde $\Delta f = f(x) - f(x_0)$ şi $\Delta x = x - x_0$.

Am numit diferenţialele lui f în x_0 aplicaţia $T: \mathbb{R} \to \mathbb{R}$, $T(h) = f'(x_0)h$. De obicei aplicaţia T se notează $df(x_0)$. Prin urmare avem $df(x_0)(h) = f'(x_0)h$. Deoarece pentru $f(x) = x$, $f'(x_0) = 1$ şi $dx(x_0)(h) = h$ pentru $h \in \mathbb{R}$ rezultă:

$$df(x_0) = f'(x_0) \cdot dx(x_0)$$

Dacă f este derivabilă în orice $x_0 \in (a,b)$ relaţia precedentă se scrie:

$$df = f'(x) \cdot dx$$

Dacă revenim la variaţia funcţiei f obţinem:

$$\Delta f \approx f'(x_0) \cdot dx = df(x_0)(\Delta x)$$

din care constatăm că diferenţiala lui f aproximează variaţia lui f. Pentru cazul mai multor variabile (de exemplu două) $z = f(x,y)$ se introduce noţiunea de diferenţială totală.

Definiţia 4.3.

Se numeşte diferenţială totală a funcţiei f expresia:

$$dz = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy$$

Pentru trei variabile expresia corespunzătoare este asemănătoare. Ca şi la funcţiile de o variabilă şi în cazul general al mai multor variabile diferenţiala totală se poate folosi la aproximarea variaţiei funcţiei atunci cînd se trece de la un punct la un punct vecin. Facem observaţia că dz este tot o funcţie de variabilele x, y şi deci iarăşi ne putem pune problema diferenţialei totale numită diferenţială totală de ordinul doi. Presupunînd existenţa derivatelor parţiale secunde şi egalitatea celor mixte avem:

$$d^2z = d(dz) = d\left(\frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy\right) = d\left(\frac{\partial f}{\partial x}\right)dx + d\left(\frac{\partial f}{\partial y}\right)dy =$$

$$= \left(\frac{\partial^2 f}{\partial x^2}dx + \frac{\partial^2 f}{\partial x \partial y}dy\right)dx + \left(\frac{\partial^2 f}{\partial y \partial x}dx + \frac{\partial^2 f}{\partial y^2}dy\right)dy =$$

$$= \frac{\partial^2 f}{\partial x^2}dx^2 + 2\frac{\partial^2 f}{\partial x \partial y}dx\,dy + \frac{\partial^2 f}{\partial y^2}dy^2$$

Inductiv putem defini $d^n z = d(d^{n-1}z)$.

Observaţie: Se convine ca f din expresia diferenţialei de ordinul I să se scrie în afara unei paranteze obţinîndu—se un binom formal care apoi se comportă chiar ca un binom la diferite puteri. Avem:

$$dz = \left(\frac{\partial}{\partial x}dx + \frac{\partial}{\partial y}dy\right)f$$

Atunci:

$$d^2z = \left(\frac{\partial}{\partial x}dx + \frac{\partial}{\partial y}dy\right)^{(2)}f$$

O formă analoagă întîlnim la 3,4,...,n variabile. Dacă (x, y) este dintr—o vecinătate a lui (a, b) atunci are loc egalitatea:

$$f(x,y) = f(a,b) + df(a,b) + \frac{1}{2!}d^2f(a,b) + \ldots + \frac{1}{n!}d^nf(a,b) + \ldots$$

cunoscută sub numele de formula lui Taylor pentru funcţii de mai multe variabile. (Fără demonstraţie).

Definiţia 4.4.

Se spune că punctul $M_0(x_0,y_0)$ este un punct de extrem pentru funcţia $z = f(x,y)$ dacă pentru o întreagă vecinătate a lui M_0 diferenţa $f(x,y) - f(x_0,y_0)$ îşi păstrează un semn constant. Dacă $f(x,y) - f(x_0,y_0) \geq 0$ extremul se numeşte punct de minim iar dacă $f(x,y) - f(x_0,y_0) \leq 0$ extremul se numeşte punct de maxim.

Teorema 4.2.

Dacă funcţia $f(x,y)$ are un maxim sau un minim în M_0 acelaşi lucru este adevărat şi pentru funcţia $f(x_0,y)$, $f(x,y_0)$.

Demonstraţie: Din $f(x,y) - f(x_0,y_0) \geq 0$ pentru orice $M(x,y)$ din vecinătatea lui $M_0(x_0,y_0)$ rezultă $f(x,y_0) - f(x_0,y_0) \geq 0$ şi analog $f(x_0,y) - f(x_0,y_0) \geq 0$ şi deci M_0 este extrem de acelaşi fel pentru funcţiile de o variabilă obţinute.

Consecinţă: În conformitate cu teorema lui Fermat rezultă:

$$\frac{\partial f}{\partial x}(x_0,y_0) = 0 \text{ şi } \frac{\partial f}{\partial y}(x_0,y_0) = 0.$$

Anularea derivatelor parţiale de ordinul I într—un punct (x_0,y_0) reprezintă condiţiile necesare dar nu şi suficiente de extrem.

Ca şi la funcţiile de o variabilă soluţiile sistemului format din anularea derivatelor parţiale de ordinul I le vom numi puncte critice.

Să notăm cu $A = \frac{\partial^2 f}{\partial x^2}$, $B = \frac{\partial^2 f}{\partial x \partial y}$, $C = \frac{\partial^2 f}{\partial y^2}$.

Teorema 4.3.

Condiţia suficientă ca un punct critic să fie punct de extrem este ca $B^2 - AC < 0$. Dacă $A<0$ $(C<0)$ punctul este de maxim iar dacă $A>0$ $(C>0)$ punctul este de minim.

Demonstraţie: Fie $M_0(x_0,y_0)$ un punct critic. Urmează că (x_0,y_0) este soluţie pentru sistemul:

$$\frac{\partial f}{\partial x} = 0 \qquad \frac{\partial f}{\partial y} = 0$$

Pentru M_0 formula lui Taylor se scrie:

$$f(x,y) = f(x_0,y_0) + df(x_0,y_0) + \frac{1}{2!}d^2f(x_0,y_0) + \ldots$$

Deci $f(x,y) - f(x_0,y_0) = df(x_0,y_0) + \frac{1}{2!}d^2f(x_0,y_0) + \ldots.$

Dar $df(x_0,y_0) = \frac{\partial f}{\partial x}(x_0,y_0)dx + \frac{\partial f}{\partial y}(x_0,y_0)dy = 0$. Deci rămîne

$$f(x,y) - f(x_0,y_0) = \frac{1}{2!}d^2f(x_0,y_0) + \frac{1}{3!}d^3f(x_0,y_0) + \ldots.$$

Semnul diferenţei va fi dat de semnul termenului $d^2f(x_0,y_0)$. Dar $d^2f(x_0,y_0) = \frac{\partial^2 f}{\partial x^2}dx^2 + 2\frac{\partial^2 f}{\partial x \partial y}dx\,dy + \frac{\partial^2 f}{\partial y^2}dy^2$ şi deci:

$$d^2f(x_0,y_0) = \frac{\partial^2 f}{\partial x^2}dx^2 + 2\frac{\partial^2 f}{\partial x \partial y}dx\,dy + \frac{\partial^2 f}{\partial y^2}dy^2$$

sau

$$d^2f(x_0,y_0) = dy^2[At^2 + 2Bt + C]$$

unde $t = \frac{dx}{dy}$.

Dacă $B^2 - AC < 0$ rezultă că trinomul din paranteză are rădăcini complexe şi deci pentru orice t are semnul lui A şi deci concluzia din teoremă.

Observaţii:

1) Dacă dădeam factor pe dx^2, rolul lui A era preluat de C.

2) A şi C au acelaşi semn deoarece $B^2 - AC < 0 \rightarrow B^2 < AC$ adică $AC > 0$.

BCU ALEXANDRU IOAN CUZA LIBRARY

3) Dacă $B^2-AC>0$ trinomul îşi schimbă semnul deci M_0 nu este extrem.

4) Dacă $B^2-AC=0$ nu putem decide (avem caz de dubiu).

Exemplu:

Fie $z=x^3+3xy-15x-12y$. Să se afle extremele lui z.

Formăm sistemul $\frac{\partial z}{\partial x}=0$, $\frac{\partial f}{\partial y}=0$. Se obţine:

$$\begin{cases} x^2+y^2-5=0 \\ xy-2=0 \end{cases}$$

cu soluţiile (1,2), (2,1), (-1,-2), (-2,-1).

Făcînd derivatele secunde şi calculînd A, B, C aflăm că primul punct nu este punct de extrem, al doilea este minim şi ultimul punct de maxim.

Evident că teorema 4.3. are loc numai pentru două variabile. Pentru mai multe variabile termenul din paranteză care decide semnul diferenţei nu mai este un trinom ci un polinom care duce la studiul unor forme pătratice. Deci dacă forma pătratică obţinută este pozitiv definită punctul critic este extrem, dacă este negativ definită este extrem iar cînd nu este definită punctul nu este de extrem.

§ 3. METODA CELOR MAI MICI PĂTRATE

În numeroase cazuri sîntem conduşi la situaţii în care trebuie să stabilim corelaţia (legitatea) între două mulţimi de date care reprezintă cauze şi efecte într-un anume cadru.

Fie datele:

x	x_1	x_2	...	x_n
y	y_1	y_2	...	y_n

Admitem că funcţia căutată este de forma $y=ax+b$ cu a, b necunoscuţi. Să notăm cu $\overline{y}_i=ax_i+b$ şi cu $\varepsilon_i=\overline{y}_i-y_i$. Numind ε_i-erorile datorate unor cauze

întîmplătoare putem spune că dorima, b aşa fel ca erorile să fie cît mai mici. Practica arată că acest lucru se realizează atunci cînd suma pătratelor erorilor este cît mai mică. Fiind pozitive suma va fi minimă cînd pătratele vor fi cele mai mici. Deci aşa se justifică numele metodei. Deci urmărim $\sum_{i=1}^{n} \varepsilon_i^2 = \min$. Înlocuind se obţine $E(a,b) = \sum_{i=1}^{n} (ax_i + b - y_i)^2 = \min$.

Formînd sistemul:

$$\frac{\partial E}{\partial a} = 0 \quad , \quad \frac{\partial E}{\partial b} = 0$$

obţinem:

$$\begin{cases} \sum_{i=1}^{n} (ax_i^2 + bx_i - x_i y_i) = 0 \\ \sum_{i=1}^{\infty} (ax_i + b - y_i) = 0 \end{cases}$$

sau

$$\begin{cases} \left(\sum_{i=1}^{n} x_i^2\right) a + \left(\sum_{i=1}^{n} x_i\right) b = \sum_{i=1}^{n} x_i y_i \\ \left(\sum_{i=1}^{n} x_i\right) a + \left(\sum_{i=1}^{n} 1\right) b = \sum_{i=1}^{n} y_i \end{cases}$$

Acest sistem ne dă a, b şi deci funcţia dorită. Dacă $y = ax^2 + bx + c$ se procedează pe o cale asemănătoare.În primul caz se mai spune că se face ajustarea datelor cu ajutorul unei drepte iar în al doilea cu ajutorul unei parabole.

Exemplu:

Să se facă ajustarea liniară a datelor:

x	0	1	2	3	4	5
y	-4	-1	2	2	6	7

BCU IASI CENTRAL UNIVERSITY LIBRARY

Vom aşeza datele pe coloane deoarece oferă avantaje:

x_i	y_i	x_i^2	$x_i y_i$
0	-4	0	0
1	-1	1	-1
2	2	4	4
3	2	9	6
4	6	16	24
5	7	25	35
15	12	55	64

În partea de jos sînt trecute sumele necesare. Sistemul va fi:

$$\begin{cases} 55a + 15b = 64 \\ 15a + 6b = 12 \end{cases}$$

cu soluţia $a = \frac{68}{35}$ $b = -\frac{20}{7}$. Deci legătura căutată este dată de $y = \frac{68}{35}x - \frac{20}{7}$.

§ 4. EXTREME LEGATE

în rezolvarea a numeroase probleme economice şi tehnice sîntem conduşi la modele matematice în care se cere găsirea unor soluţii care să minimizeze sau să maximizeze anumite funcţii de mai multe variabile în condiţiile unor restricţii suplimentare. Astfel de extreme se numesc extreme legate, extreme cu legături sau extreme condiţionate. De fapt astfel de probleme constituie obiectul domeniului de matematici aplicate denumită „Programare matematică" domeniu cu foarte multe aplicaţii în cercetarea şi planificarea economică.

Fie $z = f(x_1,...,x_n)$ definită pe $D \subseteq \mathbb{R}^n$ şi

$$(1)\quad \begin{cases} F_1(x_1,x_2,\ldots,x_n) = 0 \\ F_2(x_1,x_2,\ldots,x_n) = 0 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ F_m(x_1,x_2,\ldots,x_n) = 0 \end{cases}, \quad m < n$$

Sistemul (1) se numeşte sistem de legături sau de condiţii.

Extremele funcţiei $z = f(x_1,\ldots,x_n)$ pe mulţimea soluţiilor sistemului (1) se numesc extreme legate sau extreme condiţionate ale funcţiei f.

O metodă de a determina extremele condiţionate este cunoscută sub numele de *Metoda multiplicatorilor lui Lagrange*.

Fie $F(x_1,x_2,\ldots,x_n,\lambda_1,\lambda_2,\ldots,\lambda_m) = f(x_1,\ldots,x_n) + \lambda_1 F_1 + \ldots + \lambda_m F_m$ cu $\lambda_1,\ldots,\lambda_m$ parametri reali. Această funcţie se numeşte funcţia lui Lagrange iar $\lambda_1,\ldots,\lambda_m$ se numesc multiplicatorii lui Lagrange. Dacă $(a_1, a_2, \ldots, a_n, \alpha_1, \alpha_2, \ldots, \alpha_n)$ este punct critic liber pentru F atunci $(a_1,a_2,\ldots,a_n)$ este punct critic legat pentru f în condiţiile (1).

În felul acesta găsirea punctelor critice legate este redusă la găsirea punctelor staţionare libere (obişnuite) pentru F care se face după metoda prezentată. În concluzie problema se poate rezolva astfel:

1) Se formează F cu parametrii $\lambda_1,\lambda_2,\ldots,\lambda_m$ apoi

2) Se formează sistemul:

$$\begin{cases} \dfrac{\partial F}{\partial x_1} = 0 \\ \dfrac{\partial F}{\partial x_2} = 0 \\ \ldots\ldots\ldots\ldots \\ \dfrac{\partial F}{\partial x_n} = 0 \\ F_1 = 0 \\ \ldots\ldots\ldots\ldots \\ F_m = 0 \end{cases}$$

3) Se determină punctele critice de forma $(a_1,\ldots,a_n,\alpha_1,\ldots,\alpha_m)$ şi deci şi cele $(a_1,\ldots,a_n)$.

BCU IAŞI CENTRU VECHI LIBRARY

4) Se studiază semnul diferenţei $f(x_1,...,x_n) - f(a_1,...,a_n)$.

Exemplu:

1) Să se afle extremele legate pentru $f(x,y,z) = x^3+y^3+z^3$ în condiţiile $x^2+y^2+z^2 = 3$, $x>0$, $y>0$, $z>0$

Formăm $F(x,y,z,\lambda) = f(x,y,z) + \lambda F_1$, λ - parametru real.

$$F(x,y,z,\lambda) = x^3+y^3+z^3+\lambda(x^2+y^2+z^2-3)$$

Sistemul corespunzător este:

$$\begin{cases} 3x^2+2\lambda x = 0 \\ 3y^2+2\lambda y = 0 \\ 3z^2+2\lambda z = 0 \\ x^2+y^2+z^2-3 = 0 \end{cases}$$

Soluţia cu primele trei componente pozitive este:

$$x = y = z = 1 \qquad \lambda = -\frac{3}{2}$$

Să vedem acum semnul lui d^2F. Deoarece toate derivatele mixte sînt nule se vede că:

$$d^2F = (6x+2\lambda)dx^2+(6y+2\lambda)dy^2+(6z+2\lambda)dz^2$$

iar în punctul critic $d^2F|P = 3(dx^2+dy^2+dz^2) > 0$. Deci în $P(1,1,1)$ este un minim condiţionat.

2) Să se afle extremele legate pentru $f(x,y,z) = xy+xz+yz$ în condiţiile $xyz = 1$ $x>0$, $y>0$, $z>0$.

Avem $F(x,y,z,\lambda) = f(x,y,z) + \lambda(xyz-1)$.

Sistemul corespunzător este:

$$\begin{cases} y+z+\lambda yz = 0 \\ x+z+\lambda xz = 0 \\ x+y+\lambda xy = 0 \\ xyz = 1 \end{cases}$$

cu soluţia cu primele trei componente pozitive (1,1,1,-2).

Punctul critic $P(1,1,1)$ nu este de extrem deoarece avem:

$$d^2F = 2(1+\lambda z)dxdy + 2(1+\lambda y)dxdz + 2(1+\lambda x)dydz$$

Iar în P

$$d^2F|P = -2(dxdy + dxdz + dydz)$$

Deoarece forma pătratică $\Phi = \xi^1\xi^2 + \xi^1\xi^3 + \xi^2\xi^3$ nu este pozitiv sau negativ definită rezultă că diferenţia:

$$f(x,y,z) - f(1,1,1) \qquad (264)$$

nu îşi păstează semnul şi deci P nu este extrem.

Întrebări de autocontrol, exerciţii

1. Ce se înţelege prin vecinătate a unui punct din R^n?
2. Definiţi derivatele parţiale.
3. Ce este diferenţiala de ordinul I? Dar de ordin n?
4. Definiţi un punct de extrem pentru o funcţie de mai multe variabile.
5. Stabiliţi condiţia suficientă de extrem.
6. Enunţaţi problema ajustării şi indicaţi metoda de rezolvare.
7. În ce constă metoda multiplicatorilor lui Lagrange?
8. Aflaţi extremele funcţiilor:
   a) $z = (x-1)^2 + 2y^2$
   g) $z = x^2 + xy + y^2 + y^2 - 2x - y$
   c) $z = x^4 + y^4 - 2x^2 + 4xy - 2y^2$
   d) $z = x^2y^3(1-x-y)$

9. Folosind metoda celor mai mici pătrate să se stabilească corelaţiile dintre datele:

x	-2	0	1	3	4	5
y	1	2	3	4	5	6

BCU IAŞI UNIVERSITY LIBRARY

b)

x	1	2	3	4	5	6
y	7	1	1	2	5	6

c)

x	1	3	4	5	6	7	8
y	6	6	5	4	4	4	3

BCU IASI CENTRAL UNIVERSITY LIBRARY

EXTINDEREA NOŢIUNII DE INTEGRALĂ

§ 1. INTEGRALA RIEMANN-STIELTJES

Fie f o funcţie definită şi mărginită pe intervalul $[a,b]$ cu $a,b\in R$ şi a,b finite.

Se consideră o divizare $d=\{x_0,x_1,\ldots,x_n\}$ a intervalului $[a,b]$, adică o familie de puncte astfel ca:

$$a=x_0<x_1<\ldots<x_{n-1}<x_n=b$$

Notăm cu $\nu(d)=\max\limits_{1\le i\le n}(x_i-x_{i-1})$ şi o numim norma divizării. În fiecare interval $[x_i,x_{i-1}]$ se alege ξ_i şi se formează suma

$$\sigma_d=\sum_{i=0}^{n-1}f(\xi_i)(x_{i+1}-x_i)$$

σ_d se numeşte suma Riemann corespunzătoare divizării d şi punctelor ξ_i.

Definiţia 5.1.

Funcţia f este integrabilă (în sens Riemann) pe $[a,b]$ dacă pentru orice şir de divizări (d_n) cu $\nu(d_n)\to 0$ cînd $n\to\infty$ şi pentru orice alegere a punctelor ξ_i şirurile (σ_{d_n}) au limită comună finită. Limita I se numeşte inte-

BIBLIOTECA UNIVERSITARĂ LIBRARY

grala funcției f pe $[a,b]$ și se notează $\int_a^b f(x)dx$.

Prin urmare avem:

$$\int_a^b f(x)\,dx = \lim_{\substack{n\to\infty \\ \nu(d_n)\to 0}} \sum_{i=0}^{n-1} f(\xi_i)(x_{i+1}-x_i)$$

Așa cum s-a văzut funcțiile continue pe $[a,b]$, funcțiile monotone pe $[a,b]$, funcțiile mărginite pe și care au un număr finit de puncte de discontinuitate sînt funcții integrabile pe $[a,b]$. Există mai multe posibilități de generalizare a noțiunii de integrală. Aceste generalizări se obțin fie renunțînd la unele condiții de definiție, fie considerînd modalități de a generaliza sumele considerate, fie trecînd la mai multe dimensiuni adică la funcții de mai multe variabile.

Integrala Riemann-Stieltjes reprezintă o primă generalizare care este foarte des întîlnită.

Dacă pe intervalul $[a,b]$ se consideră pe lîngă funcția f care este mărginită și o a doua funcție g monotonă (de exemplu crescătoare) atunci divizării d i se poate asocia suma:

$$\sigma(g) = \sum_{i=0}^{n-1} f(\xi_i)[g(x_{i+1})-g(x_i)]$$

analoagă sumei Riemann și la care se reduce dacă $g(x)=x$. Procedînd la fel și cu suma Riemann se obține o generalizare a integralei Riemann numită integrala Riemann-Stieltjes și notată $\int_a^b f(x)dg(x)$. (Se citește integrala Riemann-Stieltjes a funcției f în raport cu funcția g)

Teorema 5.1.

Dacă f este continuă pe $[a,b]$ iar g monoton crescătoare continuă împreună cu derivata ei atunci:

$$\int_a^b f(x)\,dg(x) = \int_a^b f(x)\,g'(x)\,dx$$

Demonstraţie:

Considerăm suma:

$$\sigma(g) = \sum_{i=0}^{n-1} f(\xi_i)\,(g(x_{i+1}) - g(x_i))$$

Să aplicăm teorema creşterilor finite fiecărei diferenţe $g(x_{i+1}) - g(x_i)$.

Vom obţine $g(x_{i+1}) - g(x_i) = g'(\eta_i)(x_{i+1} - x_i)$, $\eta_i \in (x_i, x_{i+1})$. Luînd $\xi_i = \eta_i$ putem scrie:

$$\sigma(g) = \sum_{i=0}^{n-1} f(\eta_i)\,[g(x_{i+1}) - g(x_i)] = \sum_{i=0}^{n-1} f(\eta_i)\,g'(\eta_i)\,(x_{i+1} - x_i)$$

adică suma Riemann pentru funcţia fg'. Trecînd la limită se obţine relaţia din enunţ. De aici şi o posibilitate de calcul a acestor integrale.

Exemplu:

Fie $\int_1^2 e^x d(x^2+1) = \int_1^2 2xe^x\,dx$.

§ 2. INTEGRALE IMPROPRII

Modificînd cerinţele iniţiale impuse intervalului sau funcţiei se obţin alte extinderi.

A) Integrale cu limite infinite. Presupunem $f:[a,\infty] \to \mathbb{R}$ şi f integrabilă pe orice interval $[a,b]$.

Definiţia 5.2.

Se numeşte integrală improprie (generalizată) de specia I

$$\int_a^{\infty} f(x)\,dx = \lim_{b\to\infty} \int_a^b f(x)\,dx.$$

Dacă limita este un număr finit atunci se spune că integrala este convergentă iar în caz contrar se spune că diverge.

Exemplu:

Să se studieze integrala $\int_a^{\infty} \frac{dx}{x^{\lambda}}$ $\lambda \neq 1$.

Avem $\int_a^{\infty} \frac{dx}{x^{\lambda}} = \frac{1}{1-\lambda} x^{1-\lambda} \Big|_a^{\infty}$

Pentru $\lambda > 1$ integrala converge, pentru $\lambda < 1$ integrala diverge.

Teorema 5.2.

Dacă $\lambda > 1$ şi pentru orice x, $|f(x)| < cx^{-\lambda}$ unde c este o constantă pozitivă atunci integrala improprie $\int_a^{\infty} f(x)\,dx$ este convergentă.

Dacă $\lambda \leq 1$ şi $(x) > x^{-\lambda}$ atunci integrala improprie este divergentă.

Demonstraţie:

Din:

$$\left| \int_a^b f(x)\,dx \right| \leq \int_a^b |f(x)|\,dx \leq c \int_a^b x^{-\lambda}\,dx$$

rezultă că pentru $\lambda > 1$ $\lim_{b \to \infty} \int_a^b f(x)\,dx \leq c \frac{1}{(\lambda-1)a^{\lambda-1}}$. Deci integrala există şi limita este finită adică este convergentă.

Dacă $\lambda \leq 1$ inegalitatea $f(x) > x^{-\lambda}$ conduce la

$$\int_a^b f(x)\,dx > \int_a^b x^{-\lambda}\,dx$$

şi cum a doua integrală este divergentă urmează că şi $\int_a^{\infty} f(x)\,dx$ este tot divergentă.

BCU IAŞI CENTRAL UNIVERSITY LIBRARY

Observaţii:

a) Cazul $\int_{-\infty}^{b} f(x)\,dx$ prin schimbarea $x = -t$ se reduce imediat la cazul tratat.

b) Cazul $\int_{-\infty}^{\infty} f(x)\,dx$ se reduce la precedentul astfel. Fixăm o valoare $c \in \mathbb{R}$. Atunci putem scrie:

$$\int_{-\infty}^{\infty} f(x)\,dx = \int_{-\infty}^{c} f(x)\,dx + \int_{c}^{\infty} f(x)\,dx$$

Teorema demonstrată oferă un criteriu de convergenţă care se aplică astfel: dacă există $\lambda > 1$ astfel ca $\lim_{x\to\infty} x^{\lambda} f(x)$ să fie finită şi diferită de zero atunci $\int_{a}^{\infty} f(x)\,dx$ este convergentă. De exemplu $\int_{0}^{\infty} \frac{x^2\,dx}{x^4+x+1}$ este convergentă deoarece pentru $\lambda = 2$ $\lim_{x\to\infty} x^2 \frac{x^2}{x^4+x+1} = 1$.

Integrale din funcţii nemărginite

Fie $[a, b]$ finit dar f nemărginită în a dar mărginită în orice interval de forma $[a+\varepsilon, b]$.

Definiţa 5.3.

Prin $\int_{a}^{b} f(x)\,dx$ vom înţelege $\lim_{\varepsilon\to 0} \int_{a+\varepsilon}^{b} f(x)\,dx$. Dacă limita este finită se spune că integrala converge iar în caz contrar diverge.

Observaţii:

1) Pentru funcţii nemărginite în b integrala improprie va fi dată de:

BOUMEDIENNE UNIVERSITY LIBRARY

$$\int_a^b f(x)\,dx = \lim_{\eta \to 0} \int_a^{b-\eta} f(x)\,dx$$

Acestea se numesc integrale improprii generalizate de speţa a II-a.

2) Dacă f este nemărginită în $c \in (a,b)$ atunci descompunînd-o în sumă ajungem la situaţiile anterioare.

Exemplu. Să se stabilească convergenţa integralei:

$$\int_a^b \frac{dx}{(x-a)^\lambda} \quad \lambda \in R\,,\ \lambda \neq 1$$

Funcţia de sub integrală este nemărginită în $x = a$. Deci avem

$$\int_{a+\varepsilon}^b (x-a)^{-\lambda}\,dx = \frac{1}{1-\lambda}(x-a)^{1-\lambda}\Big|_{a+\varepsilon}^b = \frac{1}{1-\lambda}[\,(b-a)^{1-\lambda} - \varepsilon^{1-\lambda}\,]$$

Şi în acest caz rezultă că pentru $\lambda < 1$ integrala converge iar pentru $\lambda > 1$ integrala diverge. Are loc şi în acest caz o afirmaţie analoagă teoremei 5.2.

Teorema 5.3.

Dacă pentru $\lambda < 1$ şi $x > a$ are loc inegalitatea $|f(x)| \le (x-a)^{-\lambda}$ atunci integrala $\int_a^b f(x)\,dx$ este convergentă. Dacă pentru $\lambda > 1$ şi $x > a$ $f(x) > (x-a)^{-\lambda}$ atunci integrala nu este convergentă.

Demonstraţie:

Analoagă teoremei 5.2. Şi teorema 5.3. oferă un criteriu de convergenţă utilizat în aplicaţiile practice.

Să formulăm criteriul şi pentru cazul în care b este punct în care f este nemărginită. Presupunem că pentru valori apropiate de b funcţia f are forma

$f(x) = \dfrac{\varphi(x)}{(b-x)^\lambda}$ $\lambda > 0$. Dacă $\lambda < 1$ şi $\varphi(x) \le c$, c - finit atunci $\int_a^b f(x)\,dx$ converge; dacă $\lambda \ge 1$ şi $\varphi(x) \ge c > 0$ atunci integrala diverge.

BCU IASI / CENTRAL UNIVERSITY LIBRARY

§ 3. FUNCŢIILE BETA ŞI GAMA. INTEGRALA EULER-POISSON

1) Să notăm cu $\Gamma(p) = \int_0^{\infty} x^{p-1} e^{-x}\, dx$ şi să stabilim condiţii pentru p în care $\Gamma(p)$ converge.

În acest scop scriem $\Gamma(p)$ în forma:

$$\Gamma(p) = \int_0^1 x^{p-1} e^{-x}\, dx + \int_1^{\infty} x^{p-1} e^{-x}\, dx$$

Primul termen este o integrală improprie de speţa a II-a (nemărginire în 0) dacă $p < 1$. Dacă $p \geq 1$ integrala nu este improprie.

Deoarece $\lim_{x \to 0} x^{p-1} x^{1-p} e^{-x} = 1$ urmează că integrala converge dacă $1 - p < 1$, adică $p > 0$. Pentru $p < 0$ diverge. Al doilea termen este integrală cu limite infinite şi cum

$$\lim_{x \to \infty} \frac{x^{p-1} e^{-x}}{x^{-\lambda}} = \lim_{x \to \infty} \frac{x^{\lambda + p - 1}}{e^x} = 0$$

cu $\lambda > 1$ urmează că există pentru orice p.

Prin urmare $\Gamma(p)$ există pentru $p > 0$.

2) Fie $\beta(p,q) = \int_0^1 x^{p-1} (1-x)^{q-1}\, dx$. Să vedem pentru ce p, q integrala converge. Dacă $p < 1$, 0 este punct în care apare nemărginire iar dacă $q < 1$, 1 este punct în care funcţia nu este mărginită. Pentru a putea aduce clarificările necesare descompunem integrala în forma:

$$\int_0^1 x^{p-1} (1-x)^{q-1}\, dx = \int_0^{\frac{1}{2}} x^{p-1} (1-x)^{q-1}\, dx + \int_{\frac{1}{2}}^1 x^{p-1} (1-x)^{q-1}\, dx$$

Prima integrală există în condiţia $1-p<1$ adică $p>0$ iar a doua pentru $q>0$ şi deci $\beta(p,q)$ converge pentru $p,q>0$.

În cele două exemple am prezentat două integrale foarte importante numite integrale euleriene (după Euler). Să ne ocupăm mai îndeaproape de ele.

1) **Funcţia gama**, $\Gamma(p)$

Aşa cum am văzut $\Gamma(p)=\int_0^{\infty} x^{p-1}e^{-x}\,dx$ cu $p>0$. Integrînd prin părţi se obţine:

$$\Gamma(p) = (p-1)\,\Gamma(p-1)$$

sau schimbînd $\Gamma(p+1)=p\Gamma(p)$. Acestă relaţie este importantă deoarece permite micşorarea valorii argumentului pînă se ajunge la o valoare din intervalul $(0,1]$. În cazul p - număr natural evident că

$$\Gamma(n+1) = n!$$

Deci funcţia gama este o generalizare a factorialului. Pentru p diferit de întreg folosindu-se relaţia de legătură între $\Gamma(p)$ şi $\Gamma(p+1)$ şi un tabel cu valorile funcţiei cînd $p\in(0,1]$ putem calcula orice $\Gamma(p)$.

2) **Funcţia beta**

$$\beta(p,q) = \int_0^1 x^{p-1}(1-x)^{q-1}\,dx \quad p,q>0$$

Făcînd schimbarea $1-x=u$ se obţine imediat:

$$\beta(p,q) = \beta(q,p)$$

Legătura între valori ale funcţiei β obţinute prin micşorarea valorilor argumentelor are forma:

$$\beta(p+1,q+1) = \frac{pq}{(p+q)(p+q+1)}\,\beta(p,q)$$

Integrînd prin părţi obţinem:

$$\beta(p+1,q+1) = \frac{q}{p+1}\int_0^1 x^{p+1}(1-x)^{q-1}dx =$$

$$= \frac{q}{p+1}\int_0^1 [x^p + x^p(1-x)](1-x)^{q-1}dx =$$

$$= \frac{q}{p+1}\beta(p+1,q) - \frac{q}{p+1}\beta(p+1,q+1)$$

De unde:

$$\beta(p+1,q+1) = \frac{q}{p+q+1}\beta(p+1,q)$$

şi analog:

$$\beta(p+1,q+1) = \frac{p}{p+q+1}\beta(p,q+1)$$

Folosind prima obţinem: $\beta(p,q+1) = \frac{q}{p+q}\beta(p,q)$ şi apoi din a doua:

$$\beta(p+1,q+1) = \frac{pq}{(p+q)(p+q+1)}\beta(p,q)$$

Cele două funcţii sînt legate prin:

$$\beta(p,q) = \frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)}$$

Foarte multe aplicaţii în statistică şi probabilităţi are integrala: $\int_0^\infty e^{-\frac{x^2}{2}}dx = \sqrt{\frac{\pi}{2}}$ sau $\int_0^\infty e^{-x^2}dx = \frac{\sqrt{\pi}}{2}$ sau $\int_{-\infty}^{\infty} e^{-\frac{x^2}{2}}dx = \sqrt{2\pi}$. Această integrală se numeşte integrala Euler-Poisson. Se găseşte tabelată funcţia integrală a lui Lagrange.

$$\Phi(z) = \int_0^z \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}dx$$

Pentru a calcula integrala Euler-Poisson să admitem că sîntem familiarizaţi cu integrale în care intervin parametri.

Fie $I = \int_0^\infty e^{-x^2}\,dx$.

Facem schimbarea de variabilă $x = ut$, u - parametru, $u > 0$ şi obţinem $I = \int_0^\infty ue^{-u^2t^2}\,dt$. Acum să înmulţim ambii membri cu e^{-u^2}. Vom avea

$$Ie^{-u^2} = \int_0^\infty ue^{-u^2(1+t^2)}\,dt$$

Integrînd ambii membri pe $[0,\infty)$ în raport cu u obţinem:

$$I\int_0^\infty e^{-u^2}\,du = \int_0^\infty du\int_0^\infty ue^{-u^2(1+t^2)}\,dt$$

De aici schimbînd ordinea de integrare avem:

$$I^2 = \int_0^\infty dt\int_0^\infty ue^{-u^2(1+t^2)}\,du$$

Dar $\int_0^\infty ue^{-u^2(1+t^2)}\,du = -\frac{e^{-u^2(1+t^2)}}{2(1+t^2)}\Big|_0^\infty = \frac{1}{2(1+t^2)}$ şi deci:

$$I^2 = \int_0^\infty \frac{1}{2(1+t^2)}\,dt = \frac{1}{2}\,arctg\,t\Big|_0^\infty = \frac{\pi}{4}$$

şi deci $I = \frac{\sqrt{\pi}}{2}$.

Celelalte forme se obţin uşor una din alta prin schimbări potrivite de variabilă.

§ 4. INTEGRALA DUBLĂ

O altă generalizare se obţine trecîndu-se la mai multe dimensiuni. Prin trecerea la funcţii de două variabile se obţine integrala dublă.

Fie $f: D \to \mathbb{R}$ unde $D \subset \mathbb{R}^2$, D închis şi mărginit şi fie d_n şi d_m două

diviziri pe axele de coordonate:

(d_n): $a = x_0 < x_1 < \ldots < x_{n-1} < x_n = b$

$(\overline{d_m})$: $c = y_0 < y_1 < \ldots < y_{m-1} < y_m = d$

Înţelegem să acceptăm că domeniul D este mărginit de dreptunghiul $a \le x \le b \quad c \le y \le d$. Ducînd paralele la axe, domeniul D se împarte în $m \cdot n$ dreptunghiuri delimitate de $x_i \le x \le x_{i+1} \quad y_j \le y \le y_{j+1} \quad \begin{matrix} i=\overline{0,n-1} \\ j=\overline{0,m-1} \end{matrix}$.

Alegem $\xi_i \in [x_i, x_{i+1}]$, $\eta_j \in [y_j, y_{j+1}]$ şi formăm suma

$$\sigma(d_n, \overline{d_m}) = \sum_{i=0}^{n-1} \sum_{j=0}^{m-1} f(\xi_i, \eta_i)(x_{i+1} - x_i)(y_{j+1} - y_j)$$

Să notăm cu:

$$\nu_1(d_n) = \max(x_{i+1} - x_i) \ , \quad \nu_2(\overline{d_m}) = \max(y_{j+1} - y_j)$$

Definiţia 5.4.

Funcţia f se numeşte integrabilă pe D dacă pentru orice şiruri de diviziri (d_n) respectiv $(\overline{d_m})$ cu $\nu_1(d_n) \to 0$, $\nu_2(\overline{d_m}) \to 0$ atunci cînd $m, n \to \infty$ şi pentru orice alegere a perechilor (ξ_i, η_j) şirurile sumelor au limită comună finită. Limita se numeşte integrala dublă a funcţiei f pe domeniul D şi se notează $\iint\limits_{(D)} f(x,y)\,dx\,dy$.

Deci:

$$\iint\limits_{(D)} f(x,y)\,dx\,dy = \lim_{\substack{n\to\infty \\ m\to\infty \\ \nu_1 \to 0 \\ \nu_2 \to 0}} \sum_{i=0}^{n-1} \sum_{j=0}^{m-1} f(\xi_i, \eta_j)(x_{i+1} - x_i)(y_{j+1} - y_j)$$

Suma dublă din definiţie este tot suma Riemann iar D se numeşte domeniul de integrare.

Ca şi la integrala unei funcţii de o variabilă proprietăţile sumei dau proprietăţile corespunzătoare integralei duble.

Iată principalele proprietăţi:

1) Dacă $D = D_1 \cup D_2$, $D_1 \cap D_2 = \varnothing$ atunci:

$$\iint_{(D)} f(x,y)\,dx\,dy = \iint_{(D_1)} f(x,y)\,dx\,dy + \iint_{(D_2)} f(x,y)\,dx\,dy$$

2) $\iint_{(D)} dx\,dy = aria\,D.$

3) $\iint_{(D)} (f+g)\,dx\,dy = \iint_{(D)} f\,dx\,dy + \iint_{(D)} g\,dx\,dy.$

4) $\iint_{(D)} kf\,dx\,dy = k \iint_{(D)} f\,dx\,dy.$

5) Dacă $f(x,y) = f_1(x)\cdot f_2(y)$ şi $d = \{a \le x \le b\ ,\ c \le y \le d\}$

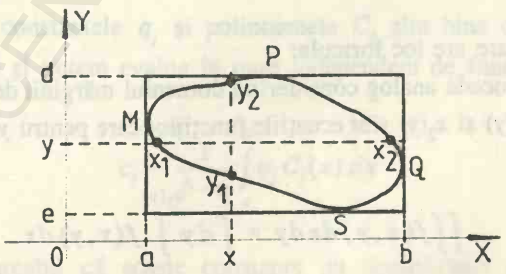
$$\iint_{(D)} f(x,y)\,dx\,dy = \int_a^b f_1\,dx \int_c^d f_2\,dy$$

6) Dacă $(\forall)\ \ (x,y) \in D\ \ f(x,y) \ge g(x,y)$ atunci:

$$\iint_{(D)} f\,dx\,dy \ge \iint_{(D)} g\,dx\,dy$$

Să vedem cum se poate calcula o integrală dublă.

Admitem că (D) este cuprins în dreptunghiul $\begin{Bmatrix} a \le x \le b \\ c \le y \le d \end{Bmatrix}$ şi orice paralelă la axe intersectează frontiera domeniului în cel mult două puncte. Grafic situaţia se prezintă astfel:



Fie y_1 şi y_2 intersecţiile unei paralele cu frontiera domeniului: Evident că acestea depind de x ales şi mai mult $y_1(x)$ reprezintă expresia funcţiei care are

drept grafic arcul de curbă MSQ iar $y_2(x)$ arcul MPQ.

Dacă admitem că $F(x) = \int_{y_1(x)}^{y_2(x)} f(x,y)$, $x \in [a,b]$ atunci are loc formula:

$$\iint_{(D)} f(x,y)\,dx\,dy = \int_a^b dx \int_{y_1(x)}^{y_2(x)} f(x,y)\,dy$$

Demonstrația se poate face pornind de la definiție. În adevăr suma Riemann se scrie:

$$\sigma(d_n, \overline{d_m}) = \sum_{i=0}^{n-1}\sum_{j=0}^{m-1} f(\xi_i, \eta_j)(x_{i+1} - x_i)(y_{j+1} - y_j) =$$

$$= \sum_{i=0}^{n-1}(x_{i+1} - x_i)\sum_{j=0}^{m-1} f(\xi_i, \eta_j)(y_{j+1} - y_j)$$

Trecînd la limită succesiv obținem:

$$\iint_{(D)} f(x,y)\,dx\,dy = \lim_{\substack{n\to\infty \\ \nu_1 \to 0}} \sum_{i=0}^{n-1}(x_{i+1} - x_i) \cdot \lim_{\substack{m\to\infty \\ \nu_2 \to 0}} \sum_{j=0}^{m-1} f(\xi_i, \eta_j)(y_{j+1} - y_j) =$$

$$= \lim_{\substack{n\to\infty \\ \nu_1 \to 0}} \sum_{i=0}^{n-1}(x_{i+1} - x_i) \int_{y_1}^{y_2} f(\xi_i, y)\,dy = \lim_{\substack{n\to\infty \\ \nu_1 \to 0}} \sum_{i=0}^{n-1} F(\xi_i)(x_{i+1} - x_i) = \int_a^b F(x)\,dx$$

Prin urmare are loc formula:

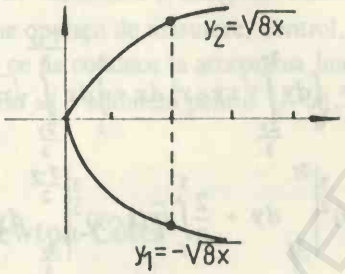
Putem proceda analog considerînd domeniul mărginit de arcele SMP și PQS. Dacă $x_1(y)$ și $x_2(y)$ sînt ecuațiile funcțiilor care pentru $y \in [c,d]$ au aceste grafice atunci:

$$\iint_{(D)} f(x,y)\,dx\,dy = \int_c^d dy \int_{x_1(y)}^{x_2(y)} f(x,y)\,dx$$

Aceste formule reduc calculul integralei duble la calculul unor integrale simple.

Exemple:

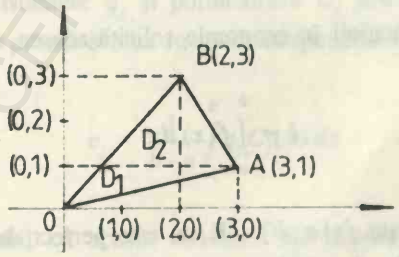
1) Să se calculeze $\iint\limits_{(D)} xy^2\,dx\,dy$ (D) fiind mărginit de $y^2 = 8x$, $x = 2$.



Avem:

$$\iint\limits_{(D)} xy^2\,dx\,dy = \int\limits_0^2 x\,dx \int\limits_{-2\sqrt{2x}}^{2\sqrt{2x}} y^2\,dy = \frac{2^{10}}{21}$$

2) Să se calculeze $\iint\limits_{(D)} \sqrt{xy - y^2}\,dx\,dy$ unde domeniul D este interiorul triunghiului cu vîrfurile $0(0,0)$, $A(3,1)$, $B(2,3)$.



Dreapta (O, A) are ecuaţia $y = \frac{x}{3}$.

Dreapta (O, B) are ecuaţia $y = \frac{3x}{2}$.

Iar dreapta (A, B) $y = -2x + 7$. Domeniul D se împarte în două D_1 , D_2 după dreapta $y = 1$ $D = D_1 \cup D_2$.

Vom avea:

$$\iint\limits_{(D)} \sqrt{xy - y^2}\, dx\, dy = \int_0^1 dy \int_{\frac{2y}{3}}^{3y} \sqrt{xy - y^2}\, dx + \int_1^3 dy \int_{\frac{2y}{3}}^{\frac{7-y}{2}} \sqrt{xy - y^2}\, dx =$$

$$= \frac{3}{2} \int_0^1 \sqrt{y}(x-y)^{\frac{3}{2}} \Big|_{\frac{2y}{3}}^{3y} dy + \frac{2}{3} \int_1^3 \sqrt{y}(x-y)^{\frac{3}{2}} \Big|_{\frac{2y}{3}}^{\frac{7-y}{2}} dy = \ldots$$

Observaţie importantă.

Procedeul utilizat la introducerea integralei duble poate fi uşor folosit şi în cazul mai multor dimensiuni. Se poate obţine astfel integrala triplă şi în general integrala multiplă.

§ 5. FORMULE DE INTEGRARE APROXIMATIVĂ

Aplicaţiile matematicii în economie solicită adesea să evaluăm cît mai exact numărul

$$I = \int_a^b f(x)\, dx$$

în condiţii în care funcţia $f: [a, b] \to \mathbb{R}$ nu este perfect determinată. În mod obişnuit ştim că funcţia f satisface anumite condiţii de regularitate ca de exemplu: „f este de clasă C_u" adică funcţia f este de u ori derivabilă şi ultima derivată este continuă, aceasta pentru valori destul de mari ale lui u.

Mai putem presupune cunoscute sau determinabile valori $y_i = f(x_i)$ pentru o submulţime finită $X = \{x_i \mid i = \overline{0,n}\} \subset [a,b]$. Desigur, cu cît mulţimea X este mai amplă, funcţia f este mai bine cunoscută şi putem spera să găsim o valoare J cît mai apropiată de valoarea I a integralei reale. De obicei însă precizarea unor valori y_i impune operaţii de măsurare, control, evaluare ce presupun cheltuieli materiale, fapt ce ne conduce la acceptarea unor compromisuri referitoare la toleranţele ce dorim să o admitem pentru $|I-J|$.

A. Formulele Newton-Côtes

Se admite aici că punctele x_i din X satisfac $x_0 = a$, $x_n = b$ şi împart $[a,b]$ în n subintervale egale. Rezultă imediat că trebuie să luăm:

$$x_k = a + k\frac{b-a}{n} \quad , \quad k = \overline{0,n}$$

Acum putem admite că funcţia f este „bine aproximată" de polinomul P de interpolare a lui Lagrange (vezi cap.II,8),

$$P(x) = \sum_{j=0}^{n} y_i q_i C_j(x)$$

Aici însă constantele q_j şi polinoamele C_j sînt bine cunoscute (după precizarea lui n) şi putem evalua în mod independent de funcţia f numerele reale.

$$c_j = \frac{1}{b-a}\int_a^b q_j C_j(x)\,dx$$

Este remarcabil că aceste constante nu depind nici de valorile a, b considerate.

Urmează imediat că vom putea lua drept valoare aproximativă pentru I numărul:

$$(NC) \quad J = (b-a)[c_0 y_0 + c_1 y_1 + \ldots + c_{n-1} y_{n-1} + c_n y_n]$$

Iată valorile coeficienţilor c_j pentru cîteva valori ale lui n:

$$n=1 \quad c_0=c_1=\frac{1}{2}$$

$$n=2 \quad c_0=c_2=\frac{1}{6}\ ,\ c_1=\frac{4}{6}$$

$$n=3 \quad c_0=c_3=\frac{1}{8}\ ,\ c_1=c_2=\frac{3}{8}$$

$$n=4 \quad c_0=c_4=\frac{7}{90}\ ,\ c_1=c_3=\frac{32}{90}\ ,\ c_4=\frac{12}{90}$$

Prezentăm mai jos calculele referitoare la $n=1$ şi $n=2$ cu scurte comentarii şi denumiri.

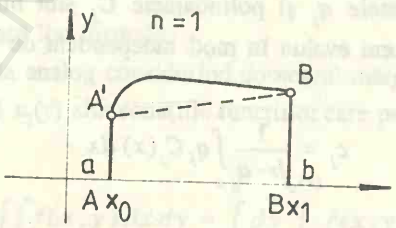
n=1. Avem $A(x)=(x-a)(x-b)x$, $C_0(x)=x-b$, $C_1(x)=x-a$, $q_0(a-b)=1=q_1(b-a)$.

$$c_0=\frac{1}{b-a}\int_a^b\frac{x-b}{a-b}\,dx=\frac{1}{(b-a)^2}\left(bx-\frac{x^2}{2}\right)\Big|_a^b=\frac{a^2-2ab+b^2}{2(b-a)^2}=\frac{1}{2}$$

Analog, $c_1=\frac{1}{2}$.

Avem deci:

$$(T)\quad J=\frac{b-a}{2}(y_1+y_2).$$



Deci aproximăm aria I a subgraficului lui f prin aria J a trapezului $ABB'A'$ prezentat în figura alăturată. Linia întreruptă $A'B'$ reprezintă aici graficul polinomului de interpolare

BCU IAŞI/CENTRAL UNIVERSITY LIBRARY

$$P(x) = y_0 \frac{x-b}{a-b} + y_1 \frac{x-a}{b-a} = \frac{y_1 - y_0}{b-a} x + \frac{y_0 b - y_1 a}{b-a}$$

Formula (T) este numită *a trapezului*.

n=2. Avem $x_1 = \frac{a+b}{2}$, $A(x) = (x-a)(x-b)(x-\frac{a+b}{2})$

$$C_0(x) = (x-b)(x-\frac{a-b}{2}) = x^2 - \frac{a+3b}{2} + \frac{b(a+b)}{2}$$

$$C_1(x) = (x-a)(x-b) = x^2 - (a+b)x + ab$$

$$C_2 = (x-a)(x-\frac{a+b}{2}) = x^2 - \frac{3a+b}{2}x + \frac{a(a+b)}{2}$$

Constantele q_i sînt date de

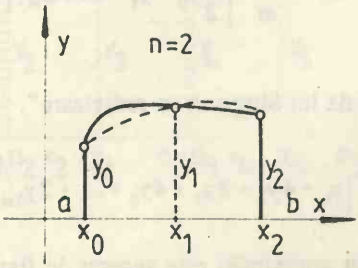
$$q_0 \frac{1}{2}(a-b)^2 = -q_1 \frac{1}{4}(a-b)^2 = q_2 \frac{1}{2}(a-b)^2 = 1$$

Urmează $c_0 = 2(b-a)^{-3} \int_a^b C_0(x)\,dx$ dar $\int_a^b C_0(x)dx = \frac{1}{12}(b-a)^3$

deci $c_0 = \frac{1}{6}$ etc.

Avem deci:

$$(S) \quad J = \frac{b-a}{6}(y_0 + 4y_1 + y_2)$$



BCU IAȘI CENTRUL UNIVERSITAR LIBRARY

Formula (S) se numeşte „a lui Simson". Interpretarea: funcţia $f:[a,b]\to\mathbb{R}$ este înlocuită printr-un arc de parabolă.

B. Formule cu redivizare

Să prezentăm intervalul $[a,b]$ drept o reuniune de intervale $[a_k, b_k]$ pentru $k=\overline{1,m}$, adică $a_1=a$, $b_1=a_2$, $b_2=a_3,\ldots,b_{m-1}=a_m$, $b_m=b$.

Notînd $I_k=\int_{a_k}^{b_k} f(x)\,dx$ vom avea desigur $I=I_1+I_2+\ldots+I_m$. Acum fiecare integrală I_k va fi aproximată printr-o formulă Newton-Côtes pentru o valoare n fixată. Cu ipoteza suplimentară că intervalele $[a_k,b_k]$ au lungimi egale formula (NC) anterioară capătă aspectul

$$J=\frac{b-a}{m}[c_0y_0+c_1y_1+\ldots,+c_{n-1}y_{n-1}+2c_my_n+$$

$$+c_1y_{n+1}+\ldots+c_{m-1}y_{2n-1}+2c_ny_{2n}+\ldots+2c_0y_{(m-1)n}+$$

$$+c_1y_{(m-1)n+1}+\ldots+c_ny_{mn}]$$

în care s-a ţinut cont şi de egalitatea $c_0=c_n$.

În particular, pentru $n=1$ se obţine „formula trapezelor".

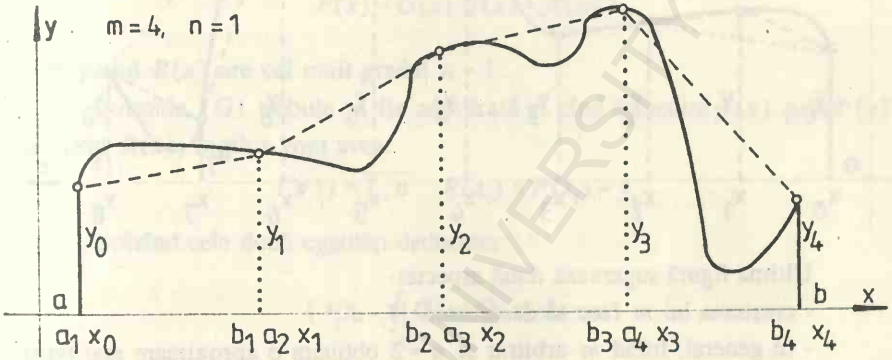
$$(T_m)\quad J=\frac{b-m}{m}\left[\frac{1}{2}y_0+y_1+\ldots+y_{m-1}+\frac{1}{2}y_m\right]$$

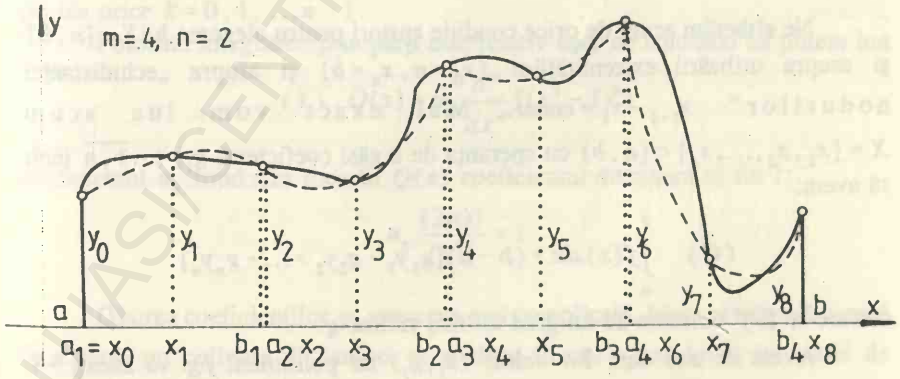
iar pentru $n=2$ „formula lui Simpson cu redivizare".

$$(S_m)\quad J=\frac{b-a}{6m}\left[y_0+4y_1+2y_2+4y_3+\ldots+2y_{2m-2}+4y_{2m-1}+y_{2m}\right]$$

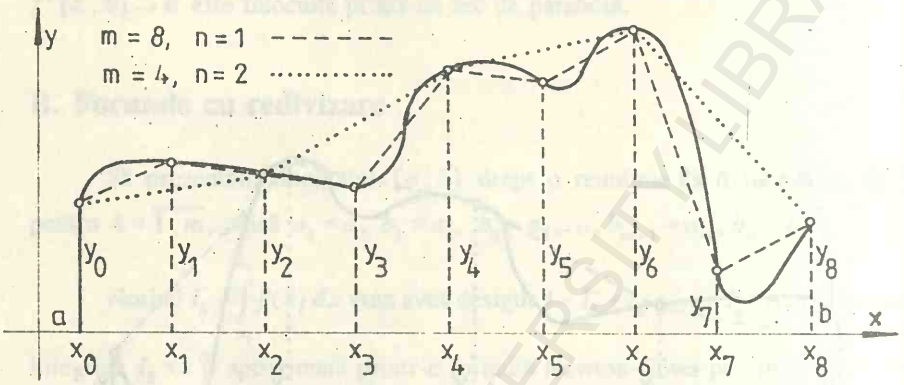
Substratul acestor aproximări este sugerat de figurile anexate.

BCU IASI CENTRAL UNIVERSITY LIBRARY





BCU CLUJ / CENTRAL UNIVERSITY LIBRARY



Ultima figură sugerează două aspecte:

- creşterea lui m face să descrească $|I - J|$
- în general, luînd m arbitrar şi $n = 2$ obţinem o aproximare mai bună decît pentru $m' = 2m$ şi $n' = 1$.

C. Formule Gauss

Ne eliberăm acum de orice condiţie apriori pentru alegerea lui $X \subset [a, b]$ şi asupra utilizării extremităţilor $(x_0 = a, x_n = b)$ şi asupra „echidistanţei nodurilor" $x_{k+1} - x_k = const$. Mai exact vom lua acum $X = \{x_1, x_2, \ldots, x_n\} \subset (a, b)$ cu speranţa de a găsi coeficienţi g_i, $i = \overline{1, n}$ încît să avem:

$$(G) \qquad \int_a^b f(x)\,dx = (b-a)[g_1 y_1 + g_2 y_2 + \ldots + g_n y_n]$$

oricare ar fi f polinom de un grad cît mai ridicat g.

Avem de ales deci $2n$ valori (x_i, g_j) iar polinomul fg va avea $g + 1$ coeficienţi arbitrari; putem spera să aibă loc $g = 2n - 1$.

Nu restrîngem generalitatea considerînd $[a, b] = [-1, +1]$.

BCU IAŞI / CENTRAL UNIVERSITY LIBRARY

Vom considera polinomul

$$G(x) = a_0(x - x_1)(x - x_2)\ldots(x - x_n)$$

Pentru un polinom $P(x)$ de grad cel mult $2n - 1$ prin algoritmul împărţirii punem în evidenţă o egalitate

$$P(x) = G(x)\,Q(x) + R(x)$$

unde restul $R(x)$ are cel mult gradul $n - 1$.

Formula (G) trebuie să fie adevărată şi cînd înlocuim $f(x)$ prin $P(x)$ sau prin $R(x)$; în plus vom avea:

$$(\forall)\, i = \overline{1, n} \quad R(x_i) = P(x_i) = y_i$$

Scăzînd cele două egalităţi deducem:

$$(*) \quad \int_{-1}^{+1} G(x)\,Q(x)\,dx = o$$

Oricare ar fi G, polinom de grad $(n - 1)$. Vom realiza $(*)$ dacă ne vom asigura că are loc

$$\int_{-1}^{+1} x^k\,Q(x)\,dx = 0$$

pentru orice $k = 0, 1, \ldots, n - 1$.

Folosind integrarea prin părţi este relativ uşor de constatat că putem lua

$$(1) \quad Q(x) = u_n \frac{d^n}{dx^n}\left[(x^2 - 1)^n\right]$$

coeficientul u_n fiind ales încît în $Q(x)$ coeficientul dominant să fie 1:

$$u_n \frac{(2n)!}{(n)!} = 1$$

Găsirea coeficienţilor g_k este ceva mai complicată. Ideea principală constă în a alege un polinom P care să se anuleze în toţi x_i cu $i \neq k$; un astfel de polinom este $P_k(x) = \dfrac{G(x)}{x - x_k}$ dar şi $\left[P_k(x)\right]^2$ sau $2\,P_k(x)\,P'_k(x)$, alegeri ce

înlesnesc calculele integralei. În final se obţine

$$g_k[(2n-1)!!](1-x_k^2)[G'(x_k)]^2 = 2(n!)^2$$

În practică se foloseşte formula lui Gauss sub forma:

$$(G_1) \quad \int_0^1 f(x)\,dx = g_1 f(x_1) + g_2 f(x_2) + \ldots + g_n f(x_n)$$

valorile lui x_i şi g_i fiind precizate în funcţie de n. Dacă $n=2$:

$$G(x) = \frac{1}{12}\frac{d^2}{dx^2}[x^2(x-1)^2] = x^2 - x + \frac{1}{6}$$

$$x_1 = 0{,}21132487$$

$$x_2 = 0{,}78867513\ ,\quad g_1 = g_2 = \frac{1}{2}$$

Întrebări de autocontrol, exerciţii

1. În ce constă generalitatea integralei Riemann-Stieltjes faţă de integrala Riemann studiată în liceu?
2. Cu motivaţi denumirea „integrală improprie"?
3. Repetaţi principalele proprietăţi ale integralei duble.
4. De ce este necesar calculul aproximativ al integralei?
5. Care sînt principalele formule de calcul aproximativ al integralelor?
6. Să se calculeze:

a) $\displaystyle\int_{-1}^{2}\frac{dx}{x}$; b) $\displaystyle\int_{-\infty}^{+\infty}\frac{dx}{x^2+4x+9}$; c) $\displaystyle\int_0^1\frac{dx}{\sqrt{1-x^2}}$;

d) $\displaystyle\int_0^{\frac{1}{2}}\frac{dx}{x\ln^2 x}$; e) $\displaystyle\int_a^{\infty}\frac{dx}{x\ln^2 x}$ $(a>1)$; f) $\displaystyle\int_0^{\infty}e^{-kx}\,dx$ $(k>0)$.

BCU IASI / CENTRAL UNIVERSITY LIBRARY

7. Să se indice limitele de integrare într-o ordine sau alta în integrala dublă $\iint\limits_{(D)} f(x,y)\,dx\,dy$ unde (D) este:

a) Dreptunghiul cu vîrfurile $O(0,0)$, $A(2,0)$, $C(2,1)$, $C(0,1)$;

b) Triunghiul de vîrfuri $O(0,0)$, $A(1,0)$, $B(1,1)$;

c) Coroana circulară limitată de cercurile de rază $r=1$, $R=2$ şi de centru comun $O(0,0)$.

8. Să se calculeze $\iint\limits_{(D)} x\,dx\,dy$ unde (D) este limitat de dreapta ce trece prin punctele $A(2,0)$, $B(0,2)$ şi arcul cercului de centru $C(0,1)$ şi $r=1$, (porţiunea care nu conţine C).

9. Să se calculeze $\iint\limits_{(D)} \frac{x\,dx\,dy}{x^2+y^2}$ unde (D) este porţiunea din parabola $y=\frac{x^2}{2}$ limitată de dreapta $y=x$.

10. Să se calculeze $\iint\limits_{(D)} xy^2\,dx\,dy$ unde (D) este limitat de $y^2=6x$ şi $x=2$.

11. Să se calculeze $\ln 2=\int\limits_0^1 \frac{dx}{1+x}$ folosind formula trapezului.

12. Să se calculeze $\ln 2$ folosind formula lui Simpson.

BCU IAŞI CENTRAL UNIVERSITY LIBRARY

ECUAŢII DIFERENŢIALE

§ 1. GENERALITĂŢI

Sînt foarte numeroase situaţiile cînd studiul unor probleme matematice, economice, biologice, mecanice etc. ne conduce la rezolvarea unor ecuaţii în care necunoscuta este o funcţie. O astfel de ecuaţie se numeşte ecuaţie funcţională. O categorie aparte de ecuaţii funcţionale o constituie acele ecuaţii funcţionale în care apar şi unele derivate ale funcţiei necunoscute, ecuaţii care se numesc ecuaţii diferenţiale.

Prezentăm în continuare două exemple.

I. Modelul matematic al creşterii populaţiei

Să considerăm populaţia unei specii, populaţie la care nu se înregistrează nici emigrări, nici imigrări. Dorim să punem în evidenţă o funcţie care, în raport cu timpul, să descrie variaţia numărului de indivizi ai populaţiei. Fie $y = f(t)$ această funcţie, ale cărei valori reprezintă numărul de indivizi din colectivitate la momentul t. Avînd în vedere interpretarea derivatei întîi a unei funcţii, urmează că derivata întîi a funcţiei f în raport cu t (notată tot $y' = f'(t)$) ne oferă viteza de creştere a populaţiei. Pe de altă parte însă, această viteză de creştere este, de fapt, tocmai diferenţa dintre rata natalităţii şi rata mortalităţii, diferenţă care se poate admite că este proporţională cu numărul indivizilor populaţiei,

adică cu $f(t)$. Prin urmare, funcţia căutată verifică o condiţie de forma:

$$y' = ky \qquad 6.1.1$$

unde k este constanta de proporţionalitate.

Aşadar, modelul matematic pe care l-am găsit este caracterizat de o ecuaţie care conţine atît funcţia necunoscută y, cît şi prima sa derivată y'.

Din (6.1.1) obţinem $\frac{dy}{dt} = ky$ sau $\frac{dy}{y} = k\,dt$, de unde obţinem mai departe:

$$\ln y = k_t + C \rightarrow y(t) = e^{k_t + C} \rightarrow y(t) = e^C e^{k_t}$$

care nu este altceva decît legea lui Malthus de creştere a populaţiei.

II. Dobînda simplă continuă

Orice proces economic antrenează numeroase fenomene, unele dintre ele putînd fi exprimate sub aspect cantitativ, valoric, prin utilizarea sumelor de bani. De exemplu, procesul de producţie dintr-o unitate economică se poate urmări cu ajutorul valorii realizate într-o anumită perioadă. Orice fenomen economic care se desfăşoară în timp şi se exprimă în bani poartă numele de operaţie financiară, cantitatea de bani care intervine în operaţia financiară respectivă numindu-se sumă. Dar este evident, pe de altă parte, că suma intrată într-o operaţie financiară variază în timpul desfăşurării, evoluţiei fenomenului economic respectiv. Variaţia unei sume de bani într-un interval de timp se numeşte *dobîndă*, operaţia financiară respectivă numindu-se operaţie de dobîndă. Dacă se notează dobînda cu D, suma iniţială cu s, iar timpul cu t, modelul matematic al reproducţiei simple ne arată că $D = k\,s\,t$, factorul de proporţionalitate k determinîndu-se imediat dacă luăm $s = 1$, $t = 1$. Obţinem deci $k = D$ = constant, adică factorul de proporţionalitate este dobînda sumei de o unitate bănească într-o unitate de timp.

În practică se ia $t = 1$ an, situaţie în care factorul de proporţionalitate se notează cu i şi se numeşte *dobîndă unitară anuală*, lucrîndu-se cu relaţia $D = i\,s\,t$.

Aplicarea practică a operaţiei de dobîndă simplă asupra sumei s este

legată de unitatea de timp, dobînda calculîndu-se la sfîrşitul perioadei. O analiză atentă a fenomenului ne arată însă că variaţia sumei nu are loc, de fapt, în mod discontinuu, şi deci D este o funcţie de timpul t : $D(t)$.

Pentru o variaţie infinitezimală a timpului obţinem deci:

$$dD(t) = ks\,dt \qquad 6.1.2$$

(unde df notează diferenţiala funcţiei f), relaţia (6.1.2) fiind o ecuaţie în care necunoscuta este funcţia $D(t)$, şi care intervine în ecuaţie prin intermediul diferenţialei sale.

Menţionînd că studiul fenomenelor de aşteptare ne va mai furniza alte exemple, putem formula:

Definiţia 6.1.1.

O *ecuaţie diferenţială* este o ecuaţie care în afara variabilelor independente şi a funcţiilor care depind de aceste variabile, mai conţin şi derivatele sau diferenţialele acestor funcţii, pînă la un anumit ordin.

Definiţia 6.1.2.

Dacă funcţiile care intervin într-o ecuaţie diferenţială depind de o singură variabilă independentă, atunci ecuaţia diferenţială se numeşte *ecuaţie diferenţială ordinară*. O ecuaţie care conţine derivatele parţiale ale funcţiilor necunoscute, în raport cu mai multe variabile independente, se numeşte *ecuaţie diferenţială cu derivate parţiale*.

În exemplele mai sus prezentate, ecuaţiile (6.1.1), (6.1.2) sînt ecuaţii diferenţiale ordinare.

În general vorbind, definiţiile 6.1.1, 6.1.2 ne arată că o ecuaţie diferenţială este o relaţie între o funcţie necunoscută şi derivatele sau diferenţialele sale pînă la un anumit ordin.

Deoarece a rezolva o ecuaţie (de orice fel) înseamnă a găsi una sau mai multe necunoscute, este evident că putem spune: „A rezolva o ecuaţie diferenţială ordinară înseamnă a găsi toate funcţiile $y = y(x)$ care satisfac o egalitate de forma:

$$\Phi\,(x; y, y', .., y^{(n)}) = 0 \qquad 6.1.3$$

Prin urmare a rezolva ecuaţia diferenţială ordinară (6.1.3) înseamnă a găsi toate funcţiile $y = y(x)$ care,împreună cu derivatele lor pînă la ordinul n, să

verifice relaţia (6.1.3) adică să transforme relaţia (6.1.3) într-o identitate în raport cu variabila independentă x.

Problema construcţiei soluţiilor unei ecuaţii diferenţiale se mai numeşte şi *integrarea ecuaţiei diferenţiale*

Definiţia 6.1.3.

Ordinul maxim la care apare (intervine) derivata într-o ecuaţie diferenţială ordinară se numeşte *ordinul ecuaţiei diferenţiale.*

§ 2. ECUAŢII DIFERENŢIALE ORDINARE DE ORDINUL ÎNTÎI. GENERALITĂŢI

Forma generală a unei ecuaţii diferenţiale ordinare de ordinul întîi este

$$\Phi(x, y, y') = 0 \qquad 6.2.1$$

sau, considerînd-o rezolvată în raport cu y':

$$y' = f(x, y) \qquad 6.2.2$$

Folosind şi notaţia $\frac{dy}{dx}$ pentru derivata de ordinul întîi a funcţiei $y = y(x)$, obţinem şi forma:

$$\frac{dy}{dx} = f(x, y) \,. \qquad 6.2.3$$

pentru o ecuaţie diferenţială ordinară de ordinul întîi.

În cazul cel mai simplu, anume acela în care $f(x, y) = f(x)$ (membrul secund din (6.2.3) nu depinde de y), ecuaţia diferenţială are deci forma $y' = f(x)$ şi determinarea soluţiilor acestei ecuaţii este - în fapt - problema fundamentală a calculului integral; mulţimea soluţiilor acestei ecuaţii este:

$$y(x) = \int f(x)\,dx + C \,,\ C \in \mathbb{R} \qquad 6.2.4$$

sau

$$y(x) = \varphi(x) + C \,,\ C \in \mathbb{R} \qquad 6.2.5$$

prin urmare, a rezolva ecuaţia diferenţială ordinară (3) înseamnă a găsi toate soluţiile $y = y(x)$ care împreună cu derivatele lor de ordinul n, să verifice relaţia (6.2.3), adică să transforme relaţia într-o identitate în raport cu variabila independentă x.

Problema construcţiei soluţiilor unei ecuaţii diferenţiale se mai numeşte şi *integrarea ecuaţiei diferenţiale.*

Astfel, în cazul cel mai simplu al ecuaţiilor diferenţiale ordinare de ordinul întîi (ca şi în general, de altfel), se poate vorbi de o familie de soluţii depinzînd de o constantă arbitrară, familie care se numeşte *integrala generală a ecuaţiei* şi pentru care se poate accepta fie o *exprimare implicită* $\Psi(x, y, C) = 0$, fie o *exprimare explicită* $\omega(x, y) = C$, $C \in \mathbb{R}$. Particularizarea constantei arbitrare C pune în evidenţă *soluţii particulare* ale ecuaţiei diferenţiale.

Interpretare geometrică. Dacă x şi y se consideră ca fiind coordonatele unui punct variabil din planul XOY, atunci atunci ecuaţia diferenţială ordinară de ordinul întîi (6.2.2) defineşte, în fiecare punct $M(x, y)$ în care funcţia $f(x, y)$ este definită (are sens), coeficientul unghiular y' al tangentei la o curbă din plan. Ca atare, soluţia $y = \varphi(x)$ a ecuaţiei diferenţiale $y' = f(x, y)$ este o curbă cu proprietatea că în fiecare punct al ei, tangenta geometrică la curbă are coeficientul unghiular definit prin $y' = f(x, y)$. Din acest motiv soluţia $y = \varphi(x)$ se mai numeşte şi *curbă integrală* a ecuaţiei diferenţiale ordinare de ordinul întîi (6.2.2). Mai mult, integrala generală $y = \varphi(x, C)$, $C \in \mathbb{R}$, reprezintă o familie de curbe integrale (o familie de curbe depinzînd de parametrul C).

Problema lui Cauchy

Am văzut mai sus că cea mai simplă ecuaţie diferenţială ordinară de ordinul întîi $y' = f(x)$ poate avea o infinitate de soluţii, dată fiind prezenţa constantei arbitrare C în expresia integralei generale. Este uşor de văzut că se obţine o soluţie complet definită pentru ecuaţia diferenţială $y' = f(x)$ dacă se *impune* o *condiţie iniţială*, adică dacă se pretinde ca funcţia căutată $y = y(x)$ să primească o valoare precizată y_0, pentru o valoare precizată $x = x_0$, adică dacă se cere ca:

$$y|_{x=x_0} = y_0 \qquad 6.2.6$$

În adevăr, integrala generală (6.2.4) se poate scrie şi sub forma

$$y(x) = \int_{x_0}^{x} f(t)\,dt + C \qquad 6.2.7$$

Şi atunci, folosind (6.2.6), obţinem $C = y_0$ deci ecuaţia diferenţială ordinară de ordinul întîi $y' = f(x)$ cu condiţia iniţială (6.2.6) are *soluţia unică*:

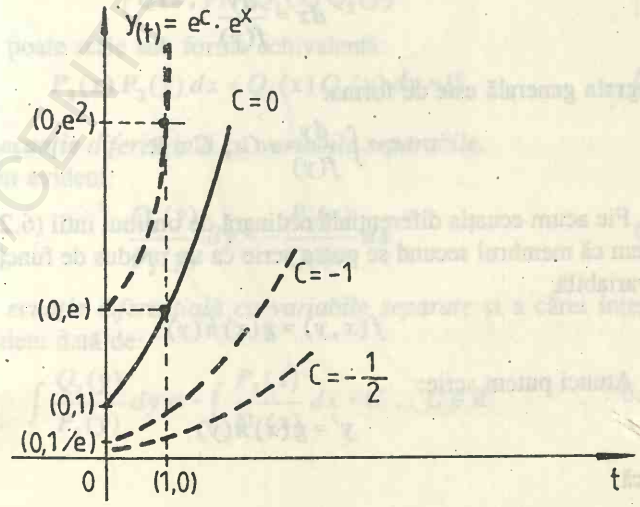
$$y(x) = \int_{x_0}^{x} f(x)\,dt + y_0 \qquad 6.2.8$$

Problema integrării unei ecuaţii difernţiale cu condiţii iniţiale poartă numele de „Problema lui Cauchy".

Referindu-ne chiar la ecuaţia diferenţială ordinară furnizată de exemplul (I) din §1, a cărei soluţie generală era:

$$y(t) = e^C e^{kt} \quad (t \geq 0)$$

dacă luăm, pentru comoditatea reprezentării grafice, $k = 1$, familia de curbe integrale $y(t) = e^C e^t$ este reprezentată, pentru anumite valori ale constantei reale C, în figura de mai jos.



BCU IAŞI/CENTRAL UNIVERSITY LIBRARY

Condiţia iniţială $y(1) = e^2$, de exemplu, ne-ar conduce la:

$$e^2 = e^C e \leftrightarrow C + 1 = 2 \leftrightarrow C = 1$$

de unde avem că prin punctul $M(1, e^2)$ trece curba integrală complet definită $y(t) = e\, e^t$ sau $y(t) = e^{t+1}$.

§ 3. TIPURI SPECIALE DE ECUAŢII DIFERENŢIALE ORDINARE DE ORDINUL ÎNTÎI

3.1. Ecuaţii cu variabile separabile

Fie ecuaţia diferenţială:

$$y' = f(y) \qquad 6.3.1$$

sau, echivalent:

$$\frac{dy}{dx} = f(y). \qquad 6.3.1'$$

Atunci, presupunînd $f(x) \neq 0$, putem scrie ecuaţia şi sub forma:

$$dx = \frac{dy}{f(y)} \qquad 6.3.2$$

şi integrala generală este de forma:

$$x = \int \frac{dy}{f(y)} + C\ ,\ C \in \mathbb{R}. \qquad 6.3.3$$

Fie acum ecuaţia diferenţială ordinară de ordinul întîi (6.2.2) în care presupunem că membrul secund se poate scrie ca un produs de funcţii de cîte o singură variabilă:

$$f(x, y) = g(x)\, h(y) \qquad 6.3.4$$

Atunci putem scrie:

$$y' = g(x)\, h(y)$$

sau încă:

$$\frac{dy}{dx} = g(x)\,h(y)$$

de unde $\frac{1}{h(x)}\,dy = g(x)\,dx$, a cărei integrală generală este deci de forma:

$$\int \frac{1}{h(y)}\,dy = \int g(x)\,dx + C\,, \quad C \in \mathbb{R} \qquad 6.3.5$$

Din cele de mai sus se vede că orice ecuaţie diferenţială ordinară de ordinul întîi de forma (6.2.2), (considerată deci rezolvată în raport cu y'), poate fi pusă sub forma:

$$P(x,y)\,dx + Q(x,y)\,dy = 0 \qquad 6.3.6$$

moment în care putem menţiona că această scriere, de fapt, nici măcar nu depinde de precizarea funcţiei necunoscute.

Definiţia 6.3.1.

Dacă funcţiile $P(x,y)$, $Q(x,y)$ se pot exprima ca produse de funcţii de cîte o singură variabilă:

$$\begin{cases} P(x,y) = P_1(x)\,P_2(y) \\ Q(x,y) = Q_1(x)\,Q_2(y) \end{cases} \qquad 6.3.7$$

atunci (6.3.6) se poate scrie sub forma echivalentă:

$$P_1(x)\,P_2(y)\,dx + Q_1(x)\,Q_2(y)\,dy = 0 \qquad 6.3.8$$

care se numeşte *ecuaţie diferenţială cu variabile separabile.*
Din (6.3.8.) avem evident:

$$\frac{Q_2(y)}{P_2(y)}\,dy = -\frac{P_1(x)}{Q_1(x)}\,dx \qquad 6.3.9$$

care se numeşte *ecuaţie diferenţială cu variabile separate* şi a cărei integrală generală este evident dată de:

$$\int \frac{Q_2(y)}{P_2(y)}\,dy = -\int \frac{P_1(x)}{Q_1(x)}\,dx + C\,, \quad C \in \mathbb{R} \qquad 6.3.10$$

Exemplu:

$$\frac{dy}{dx} = a\frac{y}{x}, \quad (a \neq 0)$$

Ecuaţia dată se mai poate scrie $\frac{dy}{y} = a\frac{dx}{x}$, adică o ecuaţie cu variabile separate. Obţinem prin integrare membru cu membru:

$$ln\,|y| = a\ln|y| + C, \quad C \in \mathbb{R}$$

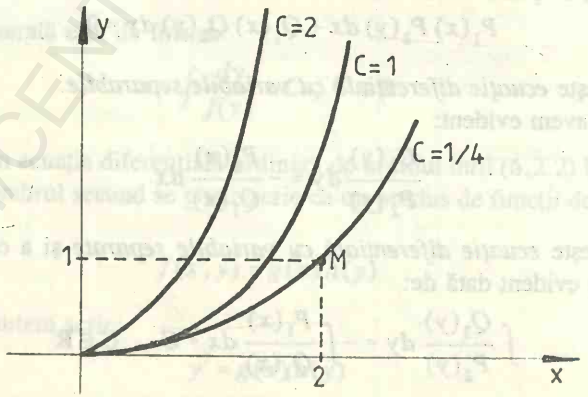
Să observăm că, din această formă, expresia care s-ar obţine pentru soluţia generală ar fi greu de manevrat. Este mai convenabil a exprima constanta de integrare tot sub forma unui logaritm natural, dintr-o constantă pozitivă arbitrară:

$$\ln|y| = a\ln|x| + \ln C, \quad C > 0$$

de unde $|y(x)| = C|x|^a$, $C > 0$, este de forma generală a soluţiei ecuaţiei diferenţiale date. Mai mult, dacă presupunem $x > 0$, $y > 0$, integrala generală are expresia:

$$y(x) = Cx^a$$

familia de curbe integrate corespunzătoare lui $a = 2$ (o familie de parabole) fiind ilustrată în figura de mai jos, pentru cîteva valori particulare ale constantei C.



BCU IASI CENTRUL DE LECTURA

Dacă s-ar fi impus şi condiţia iniţială $y(2)=1$, atunci am fi avut $1=C2^2 \Rightarrow C=\frac{1}{4}$ determinînd astfel curba integrală

$$y(x)=\frac{1}{4}x^2,$$

singura care trece prin punctul $M(2,1)$.

3.2. Ecuaţii diferenţiale ordinare omogene

Definiţia 6.3.2.

O funcţie $\varphi(x,y)$ se zice *omogenă de grad p* dacă $\varphi(tx,ty)=t^p\varphi(x,y)$.

Dacă $\varphi(x,y)$ este o funcţie omogenă de grad 0, ea se zice *omogenă* şi este evident că pentru orice funcţie omogenă are loc:

$$\varphi(x,y)=f\left(\frac{y}{x}\right) \qquad 6.3.11$$

adică se poate scrie ca o funcţie de $\frac{y}{x}$.

Definiţia 6.3.3.

O ecuaţie diferenţială ordinară de ordinul întîi se zice *omogenă* dacă este de forma:

$$y'=f\left(\frac{y}{x}\right) \qquad 6.3.12$$

Dacă păstrăm x ca variabilă independentă şi notăm (6.3.13) $u=\frac{y}{x}$, atunci evident $u=u(x)$ şi avem succesiv: $y=xu \Rightarrow y'=u+xu'$, ecuaţia (6.3.12.) devenind:

$$u+xu'=f(u)$$

BCU IASI CENTRAL UNIVERSITY LIBRARY

sau încă:

$$x u' = f(u) - u \qquad 6.3.14$$

care este o ecuaţie cu variabile separabile.

Presupunînd $f(u) - u \neq 0$, avem prin urmare:

$x \frac{du}{dx} = f(u) - u$ sau $\frac{du}{f(u) - u} = \frac{dx}{x}$, de unde, prin integrare, obţinem:

$$\ln \left| \frac{x}{C} \right| = \int \frac{du}{f(u) - u}$$

sau încă:

$$|x| = |C| \, \Psi(u) \, , \; C \in \mathbb{R} \, , \; C \neq 0 \qquad 6.3.15$$

unde $\Psi(u) = e^{\int \frac{du}{f(u)-u}}$.

Revenind la notaţiile iniţiale obţinem expresia integralei generale a ecuaţiei date:

$$|x| = |C| \, \Psi\left(\frac{y}{x}\right) \, , \; C \in \mathbb{R} \, , \; C \neq 0 \qquad 6.3.16$$

Exemplu:

Să se integreze ecuaţia diferenţială ordinară de ordin întîi:

$$y' = -\frac{x + y}{x}$$

cu condiţia iniţială $y(1) = 4$.

Avem $\varphi(x, y) = -\frac{x + y}{x} = -\left(1 + \frac{y}{x}\right) = f\left(\frac{y}{x}\right)$, deci ecuaţia dată este o ecuaţie omogenă. Notînd deci $u = \frac{y}{x}$, avem succesiv

$$y = xu \Leftrightarrow y' = u + xu' \Leftrightarrow u + xu' = -(1 + u) \Leftrightarrow$$

$$\Leftrightarrow xu' = -(1 + 2u) \Leftrightarrow x\frac{du}{dx} = -(2u + 1) \Leftrightarrow$$

$$\Leftrightarrow \frac{du}{2u + 1} = -\frac{dx}{x}$$

de unde, prin integrare, obţinem:

$$\frac{1}{2}\ln|2u+1| = \ln\left|\frac{C}{x}\right|, \; C\in\mathbb{R}, \; C>0$$

sau încă $2u+1=\left(\frac{C}{x}\right)^2$, de unde, revenind la variabilele iniţiale, obţinem integrala generală a ecuaţiei:

$$2\frac{y}{x}+1=\left(\frac{C}{x}\right)^2, \; C\in\mathbb{R}, \; C>0$$

Impunînd condiţia iniţială, obţinem:

$$2\frac{4}{1}+1=\left(\frac{C}{1}\right)^2$$

de unde $C=3$, şi deci curba integrală care trece prin punctul (1,4) este de ecuaţie:

$$2\frac{y}{x}+1=\left(\frac{3}{x}\right)^2$$

3.3. Ecuaţii diferenţiale ordinare de ordinul întîi reductibile la ecuaţii omogene

Definiţia 6.3.4.

Se numeşte ecuaţie diferenţială ordinară de ordin I *reductibilă* la o ecuaţie omogenă, o ecuaţie de forma:

$$y'=f\left(\frac{ax+by+c}{\alpha x+\beta y+\gamma}\right), \; (c,\gamma\in\mathbb{R}\setminus\{0\}) \qquad 6.3.17$$

Dacă în (6.3.17.) se fac schimbările de variabile

$$\begin{cases} x=u+m \\ y=v+n \end{cases}, \; m,n\in\mathbb{R} \qquad 6.3.18$$

BCU IAŞI CENTRAL UNIVERSITY LIBRARY

atunci ecuaţia dată se scrie sub forma:

$$v' = f\left(\frac{au + bv + am + bn + c}{\alpha u + \beta v + \alpha m + \beta n + \gamma}\right) \qquad 6.3.19$$

care, prin impunerea condiţiilor:

$$\begin{cases} am + bn + c = 0 \\ \alpha m + \beta n + \gamma = 0 \end{cases} \qquad 6.3.20$$

devine:

$$v' = f\left(\frac{au + bv}{\alpha u + \beta v}\right) \qquad 6.3.21$$

sau

$$v' = f\left(\frac{a + b\dfrac{v}{u}}{\alpha + \beta\dfrac{v}{u}}\right) \qquad 6.3.22$$

adică o ecuaţie diferenţială ordinară de ordinul întîi omogenă, cu metodă de integrare cunoscută.

Trebuie menţionat că sistemul (6.3.20) este în fapt, cel care permite „reducerea" ecuaţiei date la o ecuaţie omogenă.

Dacă sistemul (6.3.20) este compatibil determinat, adică dacă $a\beta - \alpha b \neq 0$, atunci m, n sînt unic determinaţi şi schimbarea de variabile (6.3.18) este cea care reduce într-adevăr ecuaţia dată la o ecuaţie omogenă.

Dacă însă $a\beta - \alpha b = 0$, adică dacă $\frac{a}{\alpha} = \frac{b}{\beta} = k$, atunci schimbarea de variabilă:

$$u = ax + by \qquad 6.3.23$$

transformă ecuaţia dată direct într-o ecuaţie cu variabile separabile.

Într-adevăr, din $\frac{a}{\alpha} = \frac{b}{\beta} = k$ avem şi $\frac{ax+by}{\alpha x+\beta y} = k$ sau $\alpha x + \beta y = \frac{ax + by}{k} = \frac{u}{k}$.

Folosind şi $u' = a + by' \Leftrightarrow y' = \frac{1}{b}(u' - a)$, ecuaţia devine:

$$\frac{1}{b}(u' - a) = f\left(\frac{u + c}{\frac{1}{k}u + \gamma}\right) \Leftrightarrow \frac{1}{b}(u' - a) = \varphi(u) \Leftrightarrow \tag{6.3.17}$$

$$\Leftrightarrow u' = a + b\varphi(u) \Leftrightarrow \frac{du}{dx} = a + b\varphi(u)$$

care este o ecuaţie cu variabile separabile.

Exemple:

(1) Ecuaţia diferenţială:

$$(2x + 3y - 5)y' = x - y - 5$$

se poate scrie şi sub forma $y' = \frac{x - y - 5}{2x + 3y - 5}$, adică de forma (6.3.17) cu $a = 1, b = -1, c = -5$, respectiv $\alpha = 2, \beta = 3, \gamma = -5$. Cum $a\beta - \alpha b = 5 \neq 0$. Sistemul (6.3.20) devine: $\begin{cases} m - n - 5 = 0 \\ 2m + 3n + 5 = 0 \end{cases}$ care are soluţia $\begin{cases} m = 2 \\ n = -3 \end{cases}$. Facem deci schimbarea de variabile $\begin{cases} x = u + 2 \\ y = v - 3 \end{cases}$ care transformă ecuaţia dată în $v' = \frac{u - v}{2u + 3v}$ sau încă: $v' = \frac{1 - \frac{v}{u}}{2 + 3\frac{v}{u}}$, care este o ecuaţie omogenă. Notînd $z = \frac{v}{u}$, avem pe rînd: $v = uz \Rightarrow v' = z + uz'$ şi ecuaţia omogenă devine:

$$z + uz' = \frac{1 - z}{z + 3z} \Leftrightarrow uz' = \frac{1 - z}{2 + 3z} - z \Leftrightarrow uz' = \frac{1 - 3z - 3z^2}{2 + 3z} \Leftrightarrow$$

$$\Leftrightarrow u\frac{dz}{du} = \frac{1 - 3z - 3z^2}{2 + 3z} \Leftrightarrow \frac{2 + 3z}{1 - 3z - 3z^2}dz = \frac{du}{v}$$

care este o ecuaţie cu variabile separabile.

BCU IAŞI CENTRAL UNIVERSITY LIBRARY

Prin integrare obţinem: $u = C\Psi(z)$, unde $\Psi(z) = e^{\int \frac{2+3z}{1-3z-3z^2}dz}$, sau $u = C\Psi\left(\frac{v}{u}\right)$, de unde revenind la variabilele iniţiale, obţinem:

$$x - 2 = C\Psi\left(\frac{y+3}{x-2}\right),\ C \in \mathbb{R}$$

(2) Ecuaţia diferenţială ordinară de ordinul întîi: $y' = \frac{x+2y+1}{2x+4y+3}$ are $a = 1,\ b = 2,\ c = 1,\ \alpha = 2,\ \beta = 2,\ \gamma = 3$ şi, pentru că $a\beta - \alpha b = 0$, facem notaţia: $u = x + 2y$, de unde $2x + 4y = 2(x + 2y) = 2u$ şi, din $u' = 1 + 2y'$ obţinem $y' = \frac{1}{2}(u' - 1)$, cu care ecuaţia dată devine:

$$\frac{1}{2}(u' - 1) = \frac{u+1}{2u+3} \Leftrightarrow u' - 1 = \frac{2u+2}{2u+3} \Leftrightarrow u' = \frac{4u+5}{2u+3} \Leftrightarrow$$

$$\Leftrightarrow \frac{du}{dx} = \frac{4u+5}{2u+3} \Leftrightarrow \frac{2u+3}{4u+5}du = dx$$

care este o ecuaţie cu variabile separate. Prin integrare obţinem:

$$\frac{1}{2}u + \frac{1}{8}\ln|4u+5| = x + C,\ C \in \mathbb{R}$$

Sau, revenind la variabile iniţiale:

$$\frac{1}{2}(x+2y) + \frac{1}{8}\ln|4x+8y+5| = x + C,\ C \in \mathbb{R}$$

3.4. Ecuaţia diferenţială liniară de ordinul întîi

Definiţia 6.3.5.

Se numeşte *ecuaţie diferenţială ordinară de ordinul întîi liniară* o ecuaţie diferenţială de forma:

$$y' + P(x)y = Q(x) \qquad 6.3.24$$

Dacă $Q(x) \equiv 0$, ecuaţia liniară se numeşte *omogenă*, şi *neomogenă* în caz contrar.

Să considerăm pentru început cazul ecuaţiei diferenţiale liniare de ordinul întîi omogene, adică ecuaţia:

$$y' + P(x)y = 0 \qquad 6.3.25$$

Ecuaţia (6.3.25) nu este însă decît o ecuaţie cu variabile separabile. În adevăr, din (6.3.25) avem pe rînd:

$$y' = -P(x)y \Leftrightarrow \frac{dy}{dx} = -P(x)y \Leftrightarrow \frac{dy}{y} = -P(x)\,dx$$

de unde se obţine integrala generală a ecuaţiei diferenţiale liniare de ordinul întîi omogenă:

$$y_0(x) = Ce^{-\int P'(x)dx}, \quad C \in \mathbb{R} \qquad 6.3.26$$

Soluţia generală a ecuaţiei diferenţiale liniare de ordinul întîi omogenă se mai poate scrie şi sub forma:

$$y_0(x) = Ce^{-\int_{x_0}^{x} P(t)dt}, \quad C \in \mathbb{R} \qquad 6.3.27$$

Pentru a determina soluţia generală a ecuaţiei diferenţiale liniare de ordinul întîi neomogene se foloseşte *metoda variaţiei constantelor* (Lagrange):

Se presupune, pentru moment, că $C = C(x)$ şi se obligă $y_0(x)$ - în sensul determinării lui $C(x)$ - să verifice ecuaţia diferenţială liniară de ordinul întîi neomogenă. Avem atunci:

$$y_0(x) = C(x)e^{-\int_{x_0}^{x} P(t)dt} \qquad 6.3.28$$

de unde obţinem:

$$y'_0(x) = C'(x)e^{-\int_{x_0}^{x} P(t)dt} - C(x)P(x)e^{-\int_{x_0}^{x} P(t)dt} \qquad 6.3.29$$

BCU ASSOCIATIONAL UNIVERSITY LIBRARY

şi introdusă în (6.3.24) şi ţinînd seama de faptul că $y_0(x)$ verifică (6.3.25), avem că $C = C(t)$ trebuie să verifice:

$$C'(x)e^{-\int_{x_0}^{x} P(t)dt} = Q(x) \qquad 6.3.30$$

adică o ecuaţie diferenţială ordinară de ordinul întîi cu variabile separabile. Prin integrarea ecuaţiei (6.3.30) obţinem:

$$C(x) = C_1 + \int_{x_0}^{x} Q(v)e^{\int_{x_0}^{v} P(t)dv} dt \ , \ C_1 \in \mathbb{R} \qquad 6.3.31$$

Înlocuind în (6.3.28) şi renotînd cu C constanta C_1, obţinem

$$y(x) = e^{-\int_{x_0}^{x} P(t)dt} \left(C + \int_{x_0}^{x} Q(v)e^{\int_{x_0}^{v} P(t)dt} dv \right), \ C \in \mathbb{R} \qquad 6.3.32$$

drept soluţie generală a ecuaţiei diferenţiale liniare de ordinul întîi neomogene.

Să menţionăm acum că dacă se dă condiţia iniţială

$$y(x_0) = y_o \qquad 6.3.33$$

atunci din (6.3.32) obţinem $C = y_0$ şi deci din soluţia generală (6.3.32) obţinem o soluţie complet determinată (nedepinzînd de nici un parametru):

$$y(x) = e^{-\int_{x_0}^{x} P(t)dt} \left(y_0 + \int_{x_0}^{x} Q(v)e^{\int_{x_0}^{v} P(t)dt} dv \right) \qquad 6.3.34$$

Observaţii:

(I) Peste tot în consideraţiile precedente am presupus că funcţiile $P(x), Q(x)$ sînt continue pe un interval I care conţine x_0.

(II). Fie $y_0(x)$ soluţia generală a ecuaţiei diferenţiale liniare de ordinul

întîi omogene, adică:

$$y'_0 + P(x) = 0 \qquad 6.3.35$$

şi $\bar{y}(x)$ o soluţie particulară a ecuaţiei diferenţiale liniare de ordinul întîi neomogene, adică:

$$\bar{y}' + P(x)\bar{y} = Q(x) \qquad 6.3.36$$

Din (6.3.35) şi (6.3.36) rezultă că are loc şi

$$(y_0 + \bar{y})' + P(x)(y_0 + \bar{y}) = Q(x) \qquad 6.3.37$$

adică funcţia $z(x)$ definită prin:

$$z(x) = y_0(x) + \bar{y}(x) \qquad 6.3.38$$

este soluţia generală a ecuaţiei diferenţiale liniare de ordinul întîi neomogenă.

Relaţia (6.3.38) ne arată deci că soluţia generală a ecuaţiei diferenţiale liniare de ordinul întîi neomogene se obţine din soluţia generală a ecuaţiei diferenţială liniare de ordinul întîi omogene, la care se adaugă o soluţie particulară a ecuaţiei diferenţiale de ordinul întîi neomogene.

Exemplu:

Să se integreze ecuaţia diferenţială $y' - \frac{y}{x} = x$, cu condiţia iniţială $y(-2) = 0$.

Este evident o ecuaţie diferenţială liniară de ordinul întîi neomogenă, cu $P(x) = -\frac{1}{x}$ şi $Q(x) = x$.

Folosind (6.3.32) obţinem soluţia generală:

$$y(x) = x(C + x) \ , \ C \in \mathbb{R}$$

Condiţia iniţială permiţîndu-ne să determinăm constanta C, avem $C = 2$. Soluţia căutată este dată deci de:

$$y(x) = x(2 + x)$$

BCU SIBIU CENTRAL UNIVERSITY LIBRARY

3.5. Ecuaţia lui Bernoulli

O generalizare a ecuaţiei diferenţiale liniare de ordinul întîi o constituie aşa numita ecuaţie de tip Bernoulli.

Definiţia 6.3.6.

Se numeşte ecuaţie diferenţială de ordinul întîi de tip Bernoulli o ecuaţie diferenţială de forma:

$$y' + P(x)y = Q(x)y^{\alpha}\ ,\ \alpha \in \mathbb{R} \setminus \{0,1\} \qquad 6.3.39$$

Este evident că pentru $\alpha = 0$ sau $\alpha = 1$ ecuaţia (6.3.39) se reduce la o ecuaţie diferenţială liniară de ordinul întîi sau, respectiv, la o ecuaţie diferenţială liniară de ordinul întîi omogenă (cu valorile separabile).

Ecuaţia (6.3.39) se mai poate scrie echivalent:

$$y^{-\alpha}y' + P(x)y^{1-\alpha} = Q(x) \qquad 6.3.40$$

şi, dacă se face schimbarea de variabilă:

$$u = y^{1-\alpha} \qquad 6.3.41$$

obţinem:

$$u' = (1-\alpha)y^{-\alpha}y' \qquad 6.3.42$$

cu care (6.3.40) devine:

$$\frac{1}{1-\alpha}u' + P(x)u = Q(x)$$

sau

$$u' + (1-\alpha)P(x)u = (1-\alpha)Q(x) \qquad 6.3.43$$

care este o ecuaţie diferenţială liniară de ordinul întîi neomogenă.

Exemplu:

Ecuaţia diferenţială de ordinul întîi:

BCU IASI/CENTRAL UNIVERSITY LIBRARY

$$y' + \frac{y}{x} = -xy^2$$

este evident o ecuaţie de tip Bernoulli, cu $P(x) = \frac{1}{x}$, $Q(x) = -x$, $\alpha = 2$. Facem schimbarea de variabilă indicată de (6.3.41), adică:

$$u = y^{-1} \quad sau \quad u = \frac{1}{y}$$

obţinem $u' = -\frac{y'}{y^2}$, cu care ecuaţia dată devine: $u' - \frac{u}{x} = x$, adică cea tratată în exemplul de la 3.4.

Avem deci:

$$u(x) = x(C + x) \ , \ C \in \mathbb{R}$$

şi prin urmare, revenind la variabila iniţială, obţinem integrala generală a ecuaţiei Bernoulli:

$$y(x) = \frac{1}{x(C + x)} \ , \ C \in \mathbb{R}$$

§ 4. ECUAŢII DIFERENŢIALE ORDINARE DE ORDIN SUPERIOR

Am văzut în §1 că $\Phi(x, y, y', \ldots, y^{(n)}) = 0$ este forma generală a unei ecuaţii diferenţiale ordinare de ordin n. În cele ce urmează ne vom referi, foarte pe scurt, la un caz particular al ecuaţiilor ordinare de ordin n, $n \geq 2$, şi anume la cazul liniar, cu coeficienţi constanţi, tratînd separat cazul $n = 2$ şi cazul $n \geq 3$.

4.1. Ecuaţii diferenţiale liniare de ordinul doi, cu coeficienţi constanţi

Definiţia 6.4.1.

Se numeşte ecuaţie diferenţială liniară de ordinul doi, cu coeficienţi constanţi, o ecuaţie diferenţială de forma:

BCU CENTRAL UNIVERSITY LIBRARY

$$y'' + py' + qy = f(x) \ , \ (p, q \in \mathbb{R}) \qquad 6.4.1$$

ecuaţia numindu-se *omogenă* sau *neomogenă* după cum $f(x) \equiv 0$ sau nu.

Să considerăm pentru început o ecuaţie diferenţială liniară de ordin doi, cu coeficienţi constanţi şi omogenă, adică:

$$y'' + py' + qy = 0 \ , \ (p, q \in \mathbb{R}) \qquad 6.4.2$$

şi $y_1(x), y_2(x)$ două soluţii particulare ale ecuaţiei (6.4.2).

Definiţia 6.4.2.

Soluţiile particulare $y_1(x), y_2(x)$ ale ecuaţiei diferenţiale liniare de ordinul doi, omogene, cu coeficienţi constanţi se zic *liniar independente* dacă din:

$$\alpha y_1(x) + \beta y_2(x) = 0 \ , \ (\forall) x \qquad 6.4.3$$

rezultă $\alpha = \beta = 0$.

Este uşor de văzut acum că „dacă $y_1(x)$, $y_2(x)$ sînt două soluţii liniar independente ale ecuaţiei (6.4.2), atunci şi:

$$y_0(x) = C_1 y_1(x) + C_2 y_2(x) \qquad 6.4.4$$

este soluţie pentru (6.4.2), $(\forall) C_1, C_2 \in \mathbb{R}$.

Revenind la ecuaţia (6.4.2), să presupunem că $y = e^{rx}$, $r \in \mathbb{C}$, este soluţie a acestei ecuaţii. Se obţine atunci, prin calcul direct, că r trebuie să fie soluţie a ecuaţiei:

$$r^2 + pr + q = 0 \qquad 6.4.5$$

Definiţia 6.4.3.

Ecuaţia (6.4.5) se numeşte *ecuaţia caracteristică* corespunzătoare ecuaţiei diferenţiale liniare de ordinul doi cu coeficienţi constanţi omogenie (6.4.2).

Dacă ecuaţia caracteristică (6.4.5) are rădăcini reale distincte, $r_1, r_2 \in \mathbb{R}$, $r_1 \neq r_2$, atunci $y_1 = e^{r_1 x}$ şi $y_2 = e^{r_2 x}$ sînt două soluţii liniare independente de ecuaţii (6.4.2), în sensul definiţiei 6.4.2.

Dacă ecuaţia caracteristică (6.4.5) are rădăcini complex conjugate $r_1 = \alpha + i\beta$, $r_2 = \alpha - i\beta$, $\alpha, \beta \in \mathbb{R}$, $\beta \neq 0$, atunci $y_1(x)$, $y_2(x)$ definită prin:

$$\begin{cases} y_1(x) = e^{\alpha x}\cos\beta x \\ y_2(x) = e^{\alpha x}\sin\beta x \end{cases} \qquad 6.4.6$$

sînt deasemeni două soluţii liniar independente ale ecuaţiei (6.4.2), în sensul definiţiei 6.4.2.

Să ne ocupăm, în sfîrşit, de cazul în care ecuaţia caracteristică (6.4.5) are două rădăcini reale egale: $r_1 = r_2 = -\frac{p}{2}$. Dacă presupunem că am modificat puţin coeficienţii ecuaţiei caracteristice, astfel încît ea să admită rădăcini distincte, una din ea rămînînd totuşi egală cu $-\frac{p}{2}$, (de exemplu $r_1 = -\frac{1}{2}$, $r_2 \neq r_1$) obţinem din nou două soluţii liniare independente ale ecuaţiei omogene, şi anume: $y_1 = e^{r_1 x}$, $y_2 = e^{r_2 x}$, aşa cum am văzut mai sus.

Dar atunci, conform (6.4.4) şi:

$$y(x) = \frac{e^{r_2 x} - e^{r_1 x}}{r_2 - r_1} \qquad 6.4.7$$

este soluţie a ecuaţiei diferenţiale liniare de ordinul II cu coeficienţi constanţi, modificaţi conform presupunerilor făcute. Ce se întîmplă însă dacă revenim asupra modificărilor presupuse mai sus a fi făcute asupra coeficienţilor ecuaţiei caracteristice? Natural, vom avea $r_2 \to r_1 = -\frac{p}{2}$, dar, pe o altă parte, vom avea şi că:

$$y(x) = \frac{e^{r_2 x} - e^{r_1 x}}{r_2 - r_1} \xrightarrow[r_2 \to r_1]{} x e^{r_1 x} \qquad 6.4.8$$

relaţie care ne arată că pe lîngă $y_1 = e^{r_1 x}$ este soluţie a ecuaţiei (6.4.2) şi $y_2(x) = x e^{r_1 x}$.

În adevăr, avem:

$$\begin{cases} y'_2 = e^{r_1 x} + r_1 x e^{r_1 x} \\ y''_2 = r_1 e^{r_1 x} + r_1 e^{r_1 x} + r_1^2 x e^{r_1 x} = r_1^2 x e^{r_1 x} + 2 r_1 e^{r_1 x} \end{cases} \qquad 6.4.9$$

şi înlocuind în (6.4.2) obţinem:

$$(r_1^2 x e^{r_1 x} + 2r_1 e^{r_1 x}) + p(r_1 x e^{r_1 x} + e^{r_1 x}) + q x e^{r_1 x} =$$
$$= x e^{r_1 x}(r_1^2 + p r_1 + q) + e^{r_1 x}(2r_1 + p) = 0 \ , \ \forall x$$

pentru că $r_1 = -\frac{p}{2}$ este soluţie a ecuaţiei caracteristice (6.4.5).

În concluzie, funcţie de natura şi numărul rădăcinilor reale distincte ale ecuaţiei caracteristice (6.4.5), soluţia generală $y_0(x)$ a ecuaţiei difenţiale liniare de ordinul doi, cu coeficienţi constanţi şi omogenă (6.4.2), are - potrivit şi lui (6.4.4) - următoarea formă (6.4.10):

$$\begin{cases} y_0(x) = C_1 e^{r_1 x} + C_2 e^{r_2 x} & , \text{dacă } r_1, r_2 \in \mathbb{R}, r_1 \neq r_2 \\ y_0(x) = (C_1 + C_2 x) e^{r_1 x} & , \text{dacă } r_1 = r_2 = -\frac{p}{2} \\ y_0(x) = e^{\alpha x}(C_1 \cos \beta x + C_2 \sin \beta x) & , \text{dacă } \begin{matrix} r_1 = \alpha + i\beta \\ r_2 = \alpha - i\beta \end{matrix} \ , \ \alpha, \beta \in \mathbb{R}, \beta \neq 0 \end{cases}$$

$$C_1 \ , \ C_2 \in \mathbb{R}$$

Revenind acum la ecuaţia diferenţială liniară de ordinul doi cu coeficienţi neomogenă (6.4.4) (deci cu $f(x) \neq 0$), să presupunem cunoscută o soluţie particulară $\bar{y}(x)$ a acesteia, adică:

$$\bar{y}'' + p\bar{y}' + q\bar{y} = f(x) \qquad 6.4.11$$

precum şi soluţia generală $y_0(x)$ a ecuaţiei liniare omogene corespunzătoare, adică:

$$y_0'' + p y_0' + q y_0 = 0 \qquad 6.4.12$$

Din (6.4.11) şi (6.4.12) rezultă că are loc şi

$$(y_0 + \bar{y})'' + p(y_0 + \bar{y})' + q(y_0 + \bar{y}) = f(x) \qquad 6.4.13$$

adică funcţia $y(x)$ definită prin:

$$y(x) = y_0(x) + \bar{y}(x) \qquad 6.4.14$$

este soluţia generală a ecuaţiei diferenţiale liniare de ordinul doi cu coeficienţi constanţi şi neomogene (6.4.1).

Trebuie menţionat nu numai că rezultatul l-am întîlnit deja la ecuaţiile diferenţiale de ordinul întîi, dar şi că, pe de o parte, el rămîne valabil şi pentru ecuaţii diferenţiale liniare de ordin n, $n \geq 3$, iar pe de altă parte, mai ales, că este sugerată metoda de integrare a ecuaţiilor liniare de ordin doi (şi nu numai!) cu coeficienţi constanţi şi neomogene:

Et.I - se ataşează ecuaţia omogenă corespunzătoare
- se scrie ecuaţia caracteristică şi se rezolvă
- se determină, conform (6.4.10), soluţia generală $y_0(x)$ a ecuaţiei omogene

Et.II - se determină o soluţie particulară $\bar{y}(x)$ a ecuaţiei neomogene

Et. III - se determină soluţia generală a ecuaţiei neomogene:

$$y(x) = y_0(x) + \bar{y}(x)$$

Din procedura sumar descrisă mai sus a rămas să ne ocupăm mai amănunţit de etapa a II-a. Este momentul să facem menţiunea că metoda generală de determinare a unei soluţii particulare $\bar{y}(x)$ a ecuaţiei diferenţiale liniare de ordinul doi, cu coeficienţi constanţi şi neomogene este tot metoda variaţiei constantelor (a lui Lagrange), întîlnită deja în legătură cu ecuaţiile diferenţiale de ordinul întîi, metodă aplicată soluţiei generale $y_0(x)$ a ecuaţiei omogene corespunzătoare. Aplicabilitatea metodei devine însă cu atît mai dificilă cu cît ordinul ecuaţiei diferenţiale este mai mare, pe de o parte, iar pe de altă parte, pentru forme particulare ale membrului drept $f(x)$, se poate determina mult mai uşor forma soluţiei particulare căutate $\bar{y}(x)$.

Din aceste motive ne limităm, în continuare, la a prezenta aceste cazuri particulare.

Cazul 1. Membrul drept $f(x)$ al ecuaţiei (6.4.1) are forma

$$f(x) = ae^{hx}, \quad a, h \in \mathbb{R} \qquad 6.4.15$$

Se caută atunci, prin metoda coeficienţilor nedeterminaţi soluţia particulară $\bar{y}(x)$ de forma:

6.4.16
$$\begin{cases} \bar{y}(x) = be^{hx}, (b \in \mathbb{R}), \text{ dacă } h \text{ nu este rădăcină a ecuației caracteristice, sau} \\ \bar{y}(x) = bxe^{hx}, (b \in \mathbb{R}), \text{ dacă } h \text{ este rădăcină reală simplă a ecuației caracteristice, sau} \\ \bar{y}(x) = bx^2e^{hx}, (b \in \mathbb{R}), \text{ dacă } h \text{ este rădăcină reală dublă a ecuației caracteristice} \end{cases}$$

Cazul 2. Membrul drept $f(x)$ al ecuației (6.4.1) are forma

$$f(x) = P(x)e^{hx} \qquad 6.4.17$$

unde $h \in \mathbb{R}$, iar $P(x)$ este un polinom.

Se caută atunci, prin natura coeficienților nedeterminați, soluția particulară $\bar{y}(x)$ de forma:

$$\bar{y}(x) = x^s P(x)e^{hx} \qquad 6.4.18$$

unde $\bar{P}(x)$ este un polinom, cu coeficienți nedeterminați, de același grad cu $\mathcal{P}(x)$, iar s este ordinul de multiplicitate al lui h pentru ecuația caracteristică $(s = 0,1,2)$.

Cazul 3. Membrul drept $f(x)$ al ecuației (6.4.1) are forma

$$f(x) = (a\cos lx + b\sin lx)e^{hx}, \ (a,b,l,h \in \mathbb{R}) \qquad 6.4.19$$

Se caută, prin metoda coeficienților nedeterminați, soluția particulară $\bar{y}(x)$ de forma:

6.4.20
$$\begin{cases} \bar{y}(x) = (\alpha\cos lx + \beta\sin lx)e^{hx}, \text{ dacă } h \pm il \text{ nu sînt rădăcinile ecuației caracteristice sau} \\ \bar{y}(x) = x(\alpha\cos lx + \beta\sin lx)e^{hx}, \text{ dacă } h \pm il \text{ sînt rădăcinile ecuației caracteristice} \end{cases}$$

$$(\alpha, \beta \in \mathbb{R})$$

Cazul 4. Membrul drept $f(x)$ al ecuaţiei (6.4.1) are forma:

$$f(x) = [P(x)\cos lx + Q(x)\sin lx]e^{hx} \qquad 6.4.21$$

unde $h, l \in \mathbb{R}$, iar $P(x), Q(x)$ sînt polinoame.

Se caută atunci, prin metoda coeficienţilor nedeterminaţi, soluţii particulare $\bar{y}(x)$ de forma:

$$6.4.22 \quad \begin{cases} \bar{y}(x) = [\tilde{P}(x)\cos lx + \tilde{Q}(x)\sin lx]e^{hx}, \text{ dacă } h \pm il \text{ nu sînt rădăcini} \\ \text{ale ecuaţiei caracteristice, sau} \\ \bar{y}(x) = x[\tilde{P}(x)\cos lx + \tilde{Q}(x)\sin lx]e^{hx}, \text{ dacă } h \pm il \text{ sînt rădăcini} \\ \text{ale ecuaţiei caracteristice} \end{cases}$$

unde $\tilde{P}(x), \tilde{Q}(x)$ sînt polinoame cu coeficienţi nedeterminaţi, de grad egal cu $\max\{grad\, P(x), grad\, Q(x)\}$.

Înainte de a ilustra prin cîteva exemple cele de mai sus, să facem încă o observaţie, utilă din punct de vedere practic.

Să presupunem că $f(x)$, membrul drept al ecuaţiei (6.4.1) se poate scrie sub forma

$$f(x) = f_1(x) + f_2(x) \qquad 6.4.23$$

$\bar{y}_1(x), \bar{y}_2(x)$ fiind soluţii particulare ale ecuaţiilor diferenţiale liniare de ordinul doi cu coeficienţi constanţi, neomogene, care ar avea respectiv $f_1(x), f_2(x)$ drept membri secunzi. Este uşor de constatat atunci că $\bar{y}(x) = \bar{y}_1(x) + \bar{y}_2(x)$ este o soluţie particulară a ecuaţiei neomogene cu membrul secund f dat de (6.4.23). În adevăr din:

$$y_1'' + p\bar{y}_1' + q\bar{y}_1 = f_1(x)$$
$$y_1'' + p\bar{y}_2' + q\bar{y}_2 = f_2(x)$$

obţinem prin sumare:

$$(\bar{y}_1 + \bar{y}_2)'' + p(\bar{y}_1 + \bar{y}_2)' + q(\bar{y}_1 + \bar{y}_2) = (f_1(x) + f_2(x))$$

adică rezultatul anunţat.

Exemple:

(1) $y'' - 2y' = e^{2x} + x^2 - 1$.

Ecuaţia omogenă ataşată este:

$$y'' - 2y' = 0$$

cu ecuaţia caracteristică corespunzătoare:

$r^2 - 2r = 0$, de unde $r_1 = 0$, $r_2 = 2$.

Prin urmare $y_0(x) = C_1 e^{r_1 x} + C_2 e^{r_2 x}$, $C_1, C_2 \in \mathbb{R}$, adică $y_0(x) = C_1 + C_2 e^{2x}$, $C_1, C_2 \in \mathbb{R}$ este soluţia generală a ecuaţiei omogene.

Pentru determinarea soluţiei particulare a ecuaţiei neomogene, să notăm $f_1(x) = e^{2x}$ şi $f_2(x) = x^2 - 1$, putînd astfel folosi observaţia finală, dar şi cazurile particulare prezentate.

Pentru ecuaţii neomogenă:

$$(i) \qquad y'' - 2y' = e^{2x}$$

sîntem în cazul 1, cu $h = 2$ rădăcina reală simplă a ecuaţiei caracteristice, deci căutăm $\overline{y_1}(x)$ de forma:

$$\overline{y_1}(x) = b x e^{2x}$$

Constanta $b \in \mathbb{R}$ rămîne a fi determinată din condiţia ca $\overline{y_1}(x)$ să verifice ecuaţia neomogenă (i). Avem, pe rînd:

$$\overline{y_1}' = b e^{2x} + 2 b x e^{2x}$$

$$\overline{y_1}'' = 2 b e^{2x} + 2 b e^{2x} + 4 b x e^{2x} = 4 b e^{2x} + 4 b x e^{2x}$$

şi, înlocuind în (i), obţinem:

$$4 b e^{2x} + 4 b x e^{2x} - 2(b e^{2x} + 2 b x e^{2x}) = e^{2x}$$

sau $2 b e^{2x} = e^{2x}$, de unde $b = \frac{1}{2}$, adică $\overline{y_1} = \frac{x}{2} e^{2x}$

Pentru ecuaţia neomogenă

$$(ii) \qquad y'' - 2y' = x^2 - 1$$

sîntem în cazul 2, cu $P(x) = x^2 - 1$ şi $h = 0$. Cum $h = 0$ este rădăcină reală simplă a ecuaţiei caracteristice $(s = 1)$, căutăm $\overline{y_2}(x)$ de forma

$$\overline{y_2}(x) = x(ax^2 + bx + c)$$

coeficienţi polinomului $\tilde{P}(x) = ax^2 + bx + c$ urmînd a fi determinaţi din condiţia ca $\overline{y_2}(x)$ să verifice ecuaţia neomogenă (ii). Avem, pe rînd:

$$\overline{y_2}' = 3ax^2 + 2bx + c$$

$$\overline{y_2}'' = 6ax + 2b$$

şi, înlocuind în (ii), obţinem: $6ax + 2b - 2(3ax^2 + 2bx + c) = x^2 - 1$ sau

$-6ax^2 + 2x(3a - 2b) + 2(b - c) = x^2 - 1$ de unde $\begin{cases} -6a = 1 \\ 2(3a - 2b) = 0, \\ 2(b - c) = -1 \end{cases}$

sistem care ne conduce la $\begin{cases} a = -\dfrac{1}{6} \\ b = -\dfrac{1}{4}. \\ c = \dfrac{1}{4} \end{cases}$

Aşadar $\overline{y_2}(x) = x\left(-\frac{1}{6}x^2 - \frac{1}{4}x + \frac{1}{4}\right)$ şi deci soluţia particulară a ecuaţiei neomogene va fi de forma:

$$\overline{y}(x) = \overline{y_1} + \overline{y_2} = \frac{x}{2}e^{2x} - \frac{x}{12}(2x^2 + 3x - 3)$$

de unde se obţine soluţia generală a ecuaţiei neomogene

$$y(x) = y_0(x) + \overline{y}(x) = C_1 + C_2 e^{2x} + \frac{x}{2}e^{2x} - \frac{x}{12}(2x^2 + 3x - 3), \quad C_1, C_2 \in \mathbb{R}$$

BCU IAŞI CENTRAL UNIVERSITY LIBRARY

Rezolvarea problemei lui Cauchy pentru această ecuaţie diferenţială de ordin doi cu coeficienţi constanţi presupune impunerea a două condiţii iniţiale, dat fiind faptul că integrala generală depinde de două constante arbitrare C_1, C_2. Într-o astfel de situaţie condiţiile iniţiale sînt de forma:

$$\begin{cases} y(x_0) = y_{(0)} \\ y'(x_0) = y'_{(0)} \end{cases}$$

În exemplul prezentat, în condiţiile iniţiale $y(0) = \frac{1}{8}$, $y'(0) = 1$ obţinem mai întîi:

$$y'(x) = 2C_2 e^{2x} + \frac{1}{2} e^{2x} + x e^{2x} - \frac{1}{12}(6x^2 + 6x - 3)$$

de unde:

$$\begin{cases} y(0) = \frac{1}{8} \\ y'(0) = 1 \end{cases} \Leftrightarrow \begin{cases} C_1 + C_2 = \frac{1}{8} \\ 2C_2 + \frac{1}{2} + \frac{1}{4} = 1 \end{cases} \Leftrightarrow \begin{cases} C_1 = 0 \\ C_2 = \frac{1}{8} \end{cases}$$

şi deci soluţia căutată, satisfăcînd condiţiile iniţiale impuse este:

$$y(x) = \frac{1}{8} e^{2x} + \frac{x}{2} e^{2x} - \frac{x}{12}(2x^2 + 3x - 3)$$

(2) $y'' + y = x \sin x$

Ecuaţia omogenă ataşată este: $y'' + y = 0$, cu ecuaţia caracteristică corespunzătoare $r^2 + 1 = 0$, de unde $\begin{cases} r_1 = +i \\ r_2 = -i \end{cases}$.

Prin urmare, soluţia generală a ecuaţiei omogene va fi dată de:

$$y_0(x) = C_1 \cos x + C_2 \sin x, \quad C_1, C_2 \in \mathbb{R}$$

($\alpha = 1$ şi $\beta = 1$ în (6.4.10))

Pentru determinarea soluţiei particulare $\bar{y}(x)$ a ecuaţiei neomogene să observăm că $f(x) = x \sin x$ se poate scrie sub forma

$$f(x) = (\mathcal{P}(x)\cos lx + Q(x)\sin lx)\, e^{hx}$$

cu $\mathcal{P}(x) = 0$, $Q(x) = x$, $l = 1$, $h = 0$ şi, cum $h \pm il = \pm i$ sînt exact rădăcinile ecuaţiei caracteristice, ne aflăm în cazul 4 al modalităţilor particulare de determinare a soluţiei particulare $\bar{y}(x)$. Ca atare vom alege $\bar{y}(x)$ de forma:

$$\bar{y}(x) = x\,[\tilde{\mathcal{P}}(x)\cos lx + \tilde{Q}(x)\sin lx]\, e^{hx}$$

unde evident $l = 1$, $h = 0$, iar $\tilde{\mathcal{P}}(x)$, $\tilde{Q}(x)$ sînt polinoame de grad I, adică $\tilde{\mathcal{P}}(x) = ax + b$, $\tilde{Q}(x) = cx + d$, coeficienţii a, b, c, $d \in \mathbb{R}$ rămînînd a fi determinaţi prin impunerea condiţiei ca $\bar{y}(x)$ să satisfacă ecuaţia neomogenă.

Avem aşadar: $\bar{y}(x) = x\,[(ax + b)\cos x + (cx + d)\sin x]$,

sau

$$\bar{y}(x) = (ax^2 + bx)\cos x + (cx^2 + dx)\sin x$$

de unde:

$$\bar{y}'(x) = (2ax + b)\cos x - (ax^2 + bx)\sin x + (2cx + d)\sin x +$$

$$+ (cx^2 + dx)\cos x$$

sau

$$\bar{y}'(x) = [cx^2 + x(2a + d) + b]\cos x - [ax^2 + x(b - 2c) - d]\sin x$$

şi

$$\bar{y}''(x) = (2cx + 2a + d)\cos x - [cx^2 + x(2a + d) + b]\sin x -$$

$$- (2ax + b - 2c)\sin x - [ax^2 + x(b - 2c) - d]\cos x$$

sau

$$\bar{y}''(x) = [-ax^2 + (4c - b)x + 2(a + d)]\cos x -$$

$$- [cx^2 + (4a + d)x + 2(b - c)]\sin x$$

Înlocuind în ecuaţia neomogenă, obţinem după reducerea termenilor asemenea:

$$[4cx+2(a+d)]\cos x - [4ax+2(b-c)]\sin x = x\sin x$$

de unde

$$\begin{cases} 4cx+2(a+d)=0 & ,\ \forall X \\ -4ax+2(b-c)=x & ,\ \forall x \end{cases}$$

care ne conduce la:

$$\begin{cases} c=o \\ a+d=0 \\ -4a=1 \\ b-c=0 \end{cases} \quad \text{cu soluţia} \quad \begin{cases} a=-\dfrac{1}{4} \\ b=c=0. \\ d=\dfrac{1}{4} \end{cases}$$

Prin urmare $\bar{y}(x) = -\dfrac{x^2}{4}\cos x + \dfrac{x}{4}\sin x$ şi deci soluţia generală a ecuaţiei neomogene va fi dată de:

$$y(x) = y_0(x) + \bar{y}(x) = C_1\cos x + C_2\sin x - \frac{x^2}{4}\cos x + \frac{x}{4}\sin x\ ,\quad C_1, C_2 \in \mathbb{R}$$

4.2. Ecuaţii diferenţiale liniare de ordin n $(n \geq 3)$, cu coeficienţi constanţi

Definiţia 6.4.3.

Se numeşte ecuaţie diferenţială liniară de ordin n, cu coeficienţi constanţi, o ecuaţie diferenţială de forma:

$$y^{(n)} + p_1 y^{(n-1)} + p_2 y^{(n-2)} + \ldots + p_{n-1} y' + p_n y = f(x)\ ,\quad (n \geq 1) \qquad 6.4.24$$

unde $p_1, \ldots, p_n \in \mathbb{R}$, ecuaţia numindu-se *omogenă* sau *neomogenă* după cum $f(x) \equiv 0$ sau nu.

BCU IAŞI CENTRUL UNIVERSITY LIBRARY

Se constată cu uşurinţă nu numai că (6.4.24) este o generalizare a (6.41), dar şi că toate celelalte elemente prezentate în legătură cu (6.41) se pot „transfera" în tratarea ecuaţiei (6.4.24), unele din rezultate fiind de altfel independente de ordinul ecuaţiei diferenţiale. Avem astfel, pe rînd:

Definiţia 6.4.4.

Soluţiile particulare $y_1(x)$, $y_2(x)$, ..., $y_n(x)$ ale ecuaţiei diferenţiale liniare de ordin n, cu coeficienţi constanţi şi omogene.

$$y^{(n)} + p_1 y^{(n-1)} + p_2 y^{(n-2)} + \dots p_{n-1} y' + p_n y = 0 \qquad 6.4.25$$

se zic *liniar independente* dacă din

$$\alpha_1 y_1(x) + \alpha_2 y_2(x) + \dots + \alpha_n y_n(x) = 0 \ , \ \forall x \qquad 6.4.26$$

rezultă $\alpha_1 = \alpha_2 = \dots = \alpha_n = 0$.

Este uşor de constatat că dacă $y_1(x)$, $y_2(x)$, ..., $y_n(x)$ sînt n soluţii particulare liniar independente ale ecuaţiei omogene (6.4.25), atunci:

$$y(x) = C_1 y_1(x) + C_2 y_2(x) + \dots + C_n y_n(x) \qquad 6.4.27$$

este de asemeni soluţie a ecuaţiei (6.4.25), $(\forall) C_1, C_2, \dots, C_n \in \mathbb{R}$, adică (6.4.27) ne dă forma soluţiei generale a ecuaţiei diferenţiale liniare de ordin n, cu coeficienţi constanţi şi omogenă.

Definiţia 6.4.5.

Se numeşte *ecuaţie caracteristică* (polinom caracteristic) corespunzătoare ecuaţiei diferenţiale omogene (6.4.25) ecuaţia:

$$(6.4.28) \quad r^n + p_1 r^{n-1} + p_2 r^{n-2} + \dots + p_{n-1} r + p_n = 0$$

Ca şi în cazul $n = 2$, ecuaţia caracteristică este obţinută ca urmare a impunerii condiţiei ca $g(x) = e^{rx}$ să fie o soluţie particulară a ecuaţiei omogene (6.4.25) şi, prin urmare, funcţie de natura şi ordinul de multiplicitate al rădăcinilor ecuaţiei caracteristice se construiesc soluţiile particulare liniar independente cu ajutorul cărora se obţine soluţia generală (6.4.27) a ecuaţiei omogene.

Analog celor de la cazul $n = 2$ (4.1), putem preciza acum modul de construcţie al soluţiilor particulare liniar independente ale ecuaţiei omogene (6.4.25). Astfel:

a) dacă $r = r_j$ este rădăcină reală simplă a ecuației caracteristice (6.4.28), atunci ei îi corespunde soluția particulară $y_j(x) = e^{r_j x}$

b) dacă $r = r_j$ este rădăcină reală multiplă de ordin $k, k \in \mathbb{N}^*$, a ecuației caracteristice (6.4.28), atunci ei îi corespund soluțiile particulare $y_{j_h}(x) = x^h e^{r_j x}$, $h = 0, 1, \ldots, (k-1)$.

c) dacă $r = \overline{r_j} = \alpha_j - i\beta_j$ este rădăcină complexă simplă a ecuației caracteristice ($\beta_j \neq 0$) și acestei perechi de rădăcini simple complex conjugate îi corespunde perechea de soluții particulare liniar independente:

$$\begin{cases} y_j^{(1)}(x) = e^{\alpha_j x} \cos \beta_j x \\ y_j^{(2)}(x) = e^{\alpha_j x} \sin \beta_j x \end{cases} \qquad 6.4.29$$

d) dacă $r = r_j = \alpha_j + i\beta_j$, $\alpha_j, \beta_j \in \mathbb{R}, (\beta_j \neq 0)$ este rădăcină complexă multiplă de ordin k a ecuației caracteristice, atunci și $r = \overline{r_j} = \alpha_j - i\beta_j$ este rădăcină complexă multiplă de ordin k a ecuației caracteristice, și acestei perechi de rădăcini multiple și complex conjugate îi corespund $2k$ soluții particulare liniar independente:

$$\begin{cases} y_j^{(h)_1} = x^h e^{\alpha_j x} \cos \beta_j x \quad , h = \overline{0, (k-1)} \\ y_j^{(h)_2} = x^h e^{\alpha_j x} \sin \beta_j x \quad , h = \overline{0, (k+1)} \end{cases} \qquad 6.4.30$$

În baza soluției (6.4.27) și a regulilor (6.4.29)-(6.4.32) se construiește așadar soluția generală a ecuației diferențiale liniare de ordin n, cu coeficienți constanți și omogene.

Revenind acum la ecuația neomogenă (6.4.24), dacă $\bar{y}(x)$ este o soluție particulară a acesteia, adică dacă:

$$\bar{y}^{(n)} + p_1 \bar{y}^{(n-1)} + p_2 \bar{y}^{(n-2)} + \ldots + p_{n-1} \bar{y}' + p_n \bar{y} = f(x) \qquad 6.4.31$$

iar $y_0(x)$ este soluția generală a ecuației omogene corespunzătoare (construită deci conform (6.4.27), se constată cu ușurință că:

$$y(x) = y_0(x) + \bar{y}(x) \qquad 6.4.32$$

este soluţia generală a ecuaţiei diferenţiale liniare neomogene, adică observaţiile făcute în 4.1. îşi păstrează actualitatea, în ceea ce priveşte procedura de determinare a soluţiei generale a ecuaţiei diferenţiale liniare de ordin n cu coeficienţi constanţi şi neomogene.

Soluţia particulară $\bar{y}(x)$ a ecuaţiei neomogene se determină, în general, prin aceeaşi metodă a variaţiei constantelor, dar pentru cazurile particulare ale membrului secund din (6.4.24), specificate în 4.1., $\bar{y}(x)$ se determină uşor potrivit „formei" lui $f(x)$, după regula din 4.1. adaptate cazului $n \geq 3$.

Exemplu:

Fie ecuaţia $y''' + y'' + y' + y = xe^x$, cu condiţiile iniţiale

$$y(0) = y'(0) = y'(0) = 0$$

Ecuaţia omogenă ataşată este:

$$y''' + y'' + y' + y = 0$$

cu ecuaţia caracteristică corespunzătoare:

$$r^3 + r^2 + r + 1 = 0$$

ale cărei rădăcini sînt $r_1 = -1$, $r_2 = i$, $r_3 = -i$.

Prin urmare soluţiile particulare corespunzătoare sînt $y_1(x) = e^{-x}$, $y_2(x) = \cos x$, $y_3(x) = \sin x$ şi deci $y_0(x) = C_1 y_1(x) + C_2 y_2(x) + C_3 y_3(x)$, adică

$$y_0(x) = C_1 e^{-x} + C_2 \cos x + C_3 \sin x. \quad , C_1, C_2, C_3 \in \mathbb{R}$$

este soluţia generală a ecuaţiei omogene.

Cum $f(x) = xe^x$, adică $f(x)$ este de forma $P(x)e^{hx}$, cu $\operatorname{grad} P(x) = 1$ şi $h = 1$ nu este rădăcină a ecuaţiei caracteristice, căutăm soluţia particulară $\bar{y}(x)$ de forma:

$$\bar{y}(x) = (ax + b)e^x, \text{ coeficienţii } a, b \in \mathbb{R}$$

urmînd a-i determina prin impunerea condiţiei ca $\bar{y}(x)$ să satisfacă ecuaţia neomogenă. Avem, pe rînd:

$$\bar{y}' = ae^x + (ax+b)e^x = (ax+(a+b))\,e^x$$

$$\bar{y}'' = ae^x + (ax+(a+b))e^x = (ax+(2a+b))e^x$$

$$\bar{y}''' = ae^x + (ax+(2a+b))e^x = (ax+(3a+b))e^x$$

de unde:

$$\bar{y}''' + \bar{y}'' + \bar{y}' + \bar{y} = x\,e^x$$

ne conduce la:

$$(4ax+(6a+4b))\,e^x = x\,e^x,\ \forall x \Leftrightarrow \begin{cases} 4a = 1 \\ 6a+4b = 0 \end{cases} \Leftrightarrow \begin{cases} a = \dfrac{1}{4} \\ b = -\dfrac{3}{8} \end{cases}$$

şi deci $\bar{y}(x) = (\frac{1}{4}x - \frac{3}{8})e^x$.

Prin urmare soluţia generală a ecuaţiei date este $y(x) = y_0(x) + \bar{y}(x)$, sau

$$y(x) = C_1\,e^{-x} + C_2\cos x + C_3\sin x + \frac{1}{4}(x - \frac{3}{2})e^x\ ,\quad C_1, C_2, C_3 \in \mathbb{R}$$

Pentru rezolvarea problemei lui Cauchy, avem:

$$\begin{cases} C_1 + C_2 - \dfrac{3}{8} = 0 \\ -C_1 + C_3 - \dfrac{1}{8} = 0 \\ C_1 - C_2 + \dfrac{1}{8} = 0 \end{cases} \Rightarrow \begin{cases} C_1 = \dfrac{1}{8} \\ C_2 = \dfrac{1}{4} \\ C_3 = \dfrac{1}{4} \end{cases}$$

şi deci soluţia căutată este:

$$y(x) = \frac{1}{8}e^{-x} + \frac{1}{4}(\cos x + \sin x) + \frac{1}{4}(x - \frac{3}{2})e^x$$

BCU IASI / CENTRAL UNIVERSITY LIBRARY

Întrebări de autocontrol, exerciții

1. Ce este o ecuație diferențială? Cum se clasifică ecuațiile diferențiale?
2. Ce este o soluție? Ce este soluția generală?
3. Care este problema lui Cauchy?
4. Să se rezolve următoarele ecuații diferențiale:

a) $x^2(y-1)y' + (x-1)y = 0$

b) $xyy' = 1 - x^2$

c) $xy' = xe^{\frac{y}{x}} + y + x$

d) $(2x + 3y + 1)y' + (x + 2y - 5) = 0$

e) $(2x + 3y - 1)dx + (4x + 6y - 5)dy = 0$

5) Fie $(1 + e^x)yy' = e^x$. Se cere să se afle soluția particulară care verifică condiția inițială $y = 1$ pentru $x = 0$.

6. Să se integreze ecuațiile diferențiale:

a) $y' = \frac{y}{x} - 1$

b) $y' = -\frac{x+y}{x}$

c) $\frac{dy}{dx} - \frac{y}{x} = x$

d) $\frac{dy}{dx} + \frac{2y}{x} = x^3$

7. Rezolvați:

a) $2xy\frac{dy}{dx} - y^2 + x = 0$

b) $yy' + y^2 + 4x(x+1) = 0$

c) $y'' - 5y' + 6y = 0$

d) $y'' + 4y' + 13y = 0$

e) $y'' - 4y' + 4y = x^2$

f) $y'' - y' + y = x^3 + 6$

BCU IAȘI CENTRAL UNIVERSITY LIBRARY

ELEMENTE DE CALCULUL PROBABILITĂŢILOR

§ 1. INTRODUCERE

Activitatea cotidiană a oamenilor ca şi dezvoltarea societăţii umane, în general, a fost influenţată din totdeauna de diferite fenomene (naturale, sociale, economice sau de altă natură), motiv pentru care cunoaşterea caracteristicilor şi a legilor de acţiune ale unor astfel de fenomene a preocupat şi preocupă în continuare societatea.

De regulă, analiza unui fenomen precizat presupune, sau are drept scop, delimitarea unor condiţii necesare şi suficiente pentru producerea, (apariţia), fenomenului considerat, astfel încît în final să se enunţe o afirmaţie de forma: „Dat fenomenul $\mathcal{F}$, dacă ansamblul de condiţii $\mathbb{C}$ este îndeplinit, atunci fenomenul $\mathcal{F}$ se produce cu certitudine, iar dacă nu, fenomenul $\mathcal{F}$ nu se poate produce". De exemplu, dacă ne interesează temperatura de fierbere a apei pure la presiune de o atmosferă atunci vom constata că ori de cîte ori temperatura atinge 100^0C apa începe să fiarbă, iar dacă temperatura apei este sub 100^0C apa nu fierbe.

Există fenomene pentru care precizarea condiţiilor necesare şi suficiente de producere se dovedeşte practic imposibil cel puţin la un moment dat, putîndu-se evidenţia doar o mulţime de condiţii necesare producerii fenomenului considerat, fără a avea certitudinea producereii fenomenului

atunci cînd aceste condiţii sînt îndeplinite. O astfel de situaţie se poate datora existenţei unor factori necunoscuţi, care pot influenţa producerea fenomenului, sau a unor factori despre care se ştie că influenţează producerea acestuia, dar nu se ştie exact cum se manifestă acestă influenţă.

Pentru exemplificare, să considerăm că se aruncă o monedă, fenomenul care ne interesează fiind apariţia feţei cu stema. Condiţiile C sînt condiţiile de aruncare, între care pot fi: forţa de aruncare, poziţia iniţială a monedei, natura suprafeţei pe care va cădea moneda, forma şi dimensiunile monedei, presiunea şi umiditatea aerului etc. Există însă şi alţi factori care pot influenţa apariţia uneia sau celeilalte feţe a monedei, cum ar fi: intensitatea curenţilor de aer existenţi în momentul aruncării, distribuţia masei monedei şi, eventual, alţi factori necunoscuţi sau neprecizaţi.

Pe de altă parte, chiar dacă se cunosc condiţiile necesare şi suficiente de producere ale unui fenomen dat, din punct de vedere practic, ne poate interesa dacă acest fenomen se poate produce la un moment t, cunoscîndu-se doar condiţiile la un moment t_0, anterior acestuia, îndeplinirea sau neîndeplinirea condiţiilor care să asigure producerea fenomenului la momentul t fiind determinată pe de o parte de condiţiile existente la momentul t_0, iar pe de altă parte, de o serie de factori mai mult sau mai puţin controlabili, care modifică aceste condiţii în intervalul de timp pînă la momentul t. Un exemplu edificator pentru o astfel de situaţie este mişcare browniană, sau modul în care se face prognoza timpului.

Dacă vom încerca să stabilim, cu mijloace matematice, în ce măsură un fenomen dat se produce atunci cînd sînt îndeplinite anumite condiţii, cu precizarea că aceste condiţii nu ne asigură nici de producerea, nici de neproducerea certă a fenomenului considerat, calculul probabilităţilor se dovedeşte în majoritatea cazurilor un instrument de lucru deosebit de util. Ca disciplină matematică, teoria probabilităţilor are drept obiect de studiu legile care guvernează comportarea fenomenelor aleatoare, prin fenomene aleatoare înţelegîndu-se acele fenomene a căror producere (realizare) este imprevizibilă datorită acţiunii unor factori necontrolabili sau necunoscuţi. Teoria probabilităţilor este un instrument de lucru indispensabil specialiştilor din domenii de activitate variate cum sînt: controlul de calitate, biologie, teoria erorilor experimentale, sociologie şi psihologie, medicină, economie.

§ 2. EVENIMENTE. OPERAŢII CU EVENIMENTE

În cadrul teoriei probabilităţilor noţiunea de eveniment se consideră ca noţiune primară. Avînd în vedere aplicaţiile acestei teorii la rezolvarea unor probleme practice concrete, este util să construim un suport intuitiv pentru noţiunea de eveniment. Rezumînd consideraţiile din prima parte a introducerii putem spune că, dacă se consideră un fenomen $\mathcal{F}$ şi la un moment dat t_0 este îndeplinit un ansamblu de condiţii $\mathbb{C}$, producerea (realizarea) acestui fenomen la momentul t_0 nu este certă, în general. Prin convenţie, asociem unei perechi $(\mathbb{C}, \mathcal{F})$, formată dintr-un ansamblu de condiţii $\mathbb{C}$ şi un fenomen $\mathcal{F}$, (care se poate produce sau nu atunci cînd ansamblul de condiţii $\mathbb{C}$ este îndeplinit), noţiunea de *eveniment*.

Vom spune că evenimentul asociat unei perechi $(\mathbb{C}, \mathcal{F})$, definită ca mai sus, se realizează dacă fenomenul $\mathcal{F}$ se produce atunci cînd sînt îndeplinite condiţiile $\mathbb{C}$ şi vom spune că evenimentul asociat perechii $(\mathbb{C}, \mathcal{F})$ nu se realizează, dacă fenomenul $\mathcal{F}$ nu se produce, deşi condiţiile $\mathbb{C}$ sînt îndeplinite.

În general, vom nota evenimentele cu litere mari ale alfabetului, A, B, C, ..., eventual cu indici, $A_1, A_2, \ldots$. Pentru cele ce urmează, este util ca încă de acum să distingem două evenimente speciale, unul numit *evenimentul sigur*, iar cel de al doilea *evenimentul imposibil*. Dacă perechea $(\mathbb{C}, \mathcal{F})$ este astfel construită încît îndeplinirea condiţiilor $\mathbb{C}$ la un moment dat asigură producerea fenomenului $\mathcal{F}$ evenimentul corespunzător acestei perechi îl vom numi *eveniment sigur*. Altfel spus, prin eveniment sigur înţelegem un eveniment care se produce cu certitudine.

Vom nota evenimentul sigur prin Ω. Dacă dimpotrivă, perechea $(\mathbb{C}, \mathcal{F})$ este astfel încît îndeplinirea condiţiilor $\mathbb{C}$ la un moment dat ne asigură că fenomenul $\mathcal{F}$ nu se poate produce, evenimentul corespunzător acestei perechi îl vom numi *evenimentul imposibil*. Cu alte cuvinte, prin eveniment imposibil se înţelege un eveniment care, cu certitudine, nu se produce. Evenimentul imposibil îl vom nota cu Φ.

Exemplul 7.2.1.

Să presupunem că un ansamblu de condiţii $\mathbb{C}$ este îndeplinit atunci cînd pe o anumită suprafaţă se aruncă două zaruri cu feţele marcate, ca de obicei, cu puncte de la 1 la 6.

Fie acum $\mathscr{F}_1$ fenomenul care se produce atunci cînd suma punctelor de pe cele două feţe ale zarurilor, care apar în urma aruncării, este cel mai mult 12, iar $\mathscr{F}_2$ fenomenul care se produce atunci cînd suma punctelor este cel puţin 13. Este atunci clar că perechii $(\mathbb{C}, \mathscr{F}_1)$ îi corespunde evenimentul sigur, iar perechii $(\mathbb{C}, \mathscr{F}_2)$ evenimentul imposibil, deoarece cea mai mare sumă a punctelor se obţine atunci cînd pe ambele zaruri apare faţa marcată cu 6 puncte şi, prin urmare, îndeplinirea condiţiilor $\mathbb{C}$ asigură producerea fenomenului $\mathscr{F}_2$.

Prin experienţă, sau experiment, vom înţelege observarea producerii unor fenomene atunci cînd sînt îndeplinite anumite condiţii, numite condiţii de efectuare ale experienţei. În acest mod fiecărei experienţe (reale sau imaginare) i se poate asocia un ansamblu de evenimente, fiecare eveniment corespunzînd unui rezultat posibil al experienţei, adică unui fenomen a cărui producere poate fi observată cînd experienţa are loc. Să observăm că în această situaţie fiecare eveniment din ansamblul de evenimente asociat experienţei corespunde unei perechi $(\mathbb{C}, \mathscr{F})$ în care $\mathbb{C}$ sînt condiţiile precizate de efectuare a experienţei, iar $\mathscr{F}$ un fenomen, (rezultat), care poate fi observat atunci cînd experienţa are loc.

În cele ce urmează vom considera numai evenimente dintr-un ansamblu de evenimente corespunzător unei experienţe, adică evenimente asociate unei perechi de forma $(\mathbb{C}, \mathscr{F})$ în care condiţiile $\mathbb{C}$ sînt aceleaşi pentru toate perechile considerate.

Între evenimetele dintr-un ansamblu de evenimente asociat unei experienţe pot fi stabilite unele relaţii între care menţionăm acum *implicaţia* şi *egalitatea* evenimentelor.

Definiţia 7.2.1.

Dacă se dau două evenimente, A, B, spunem că evenimentul A implică evenimentul B şi scriem $A \subset B$, dacă ori de cîte ori se produce evenimentul A se produce şi evenimetul B.

Exemplul 7.2.2.

Să presupunem că se aruncă două zaruri ale căror feţe sînt marcate, ca de obicei, cu un număr de puncte cuprins între 1 şi 6. Fie A evenimentul care se produce atunci cînd, după aruncarea zarurilor, suma punctelor este cel mult 4, iar B evenimetul care se produce cînd suma punctelor este cel mult 8. Conform definiţiei de mai sus, putem spune că evenimentul A implică evenimentul B, $A \subset B$, deoarece dacă se produce evenimentul A, (suma punctelor de pe cele două zaruri este cel mult 4), atunci se produce şi evenimentul B, (suma punctelor de pe cele două zaruri este cel mult 8).

Definiţia 7.2.2.

Fie A şi B două evenimente; spunem că evenimentele A şi B sînt egale, şi scriem $A = B$, dacă $A \subset B$ şi $B \subset A$.

Cu alte cuvinte, evenimentele A, B sînt egale atunci cînd ori de cîte ori se produce evenimentul A se produce şi evenimentul B, şi invers: ori de cîte ori se produce B se produce şi A.

Operaţii cu evenimente

Dacă se consideră un ansamblu de evenimente, acestea pot fi compuse în diferite moduri obţinîndu-se alte evenimente. Vom defini în acest paragraf *reuniunea* şi *intersecţia* a două evenimente şi *contrarul* unui eveniment.

Reuniunea a două evenimente

Definiţia 7.2.3.

Fie A, B două evenimente; prin reuniunea acestora înţelegem un nou eveniment, care se produce atunci cînd se produce măcar unul dintre evenimentele A, B.

Reuniunea evenimentelor A, B o notăm cu $A \cup B$, (citim evenimentul A reunit cu B sau evenimentul A sau B).

Exemplul 7.2.5.

Fie A, B evenimentele descrise în exemplul 7.2.3 de mai sus. Evenimentul $A \cap B$ este atunci evenimentul imposibil, $A \cap B = \Phi$, deoarece nu este posibil ca în urma aruncării zarului să apară simultan faţa marcată cu 1 şi faţa marcată cu 2.

Exemplul 7.2.6.

Să considerăm evenimentele A, B descrise în exemplul 7.2.4. Evenimentul $A \cap B$ este evenimentul care se produce atunci cînd, persoana sosită în staule va aştepta cel mult 3 minute, adică $A \cap B = A$.

Ţinînd cont de definiţia intersecţiei a două evenimente se poate verifica direct că au loc următoarele proprietăţi simple:

a) $A \cap A = A$, $A \cap B = B \cap A$,

b) $(A \cap B) \cap C = A \cap (B \cap C)$,

c) $A \cap \Omega = A$, $A \cap \Phi = \Phi$,

d) $A \cap (B \cap C) = (A \cap B) \cup (A \cap C)$,

e) $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$

în care A, B, C sînt evenimete oarecare, iar Ω, Φ sînt respectiv evenimentul sigur şi evenimentul imposibil. În cazul în care evenimentele A, B sînt astfel încît $A \cap B = \Phi$ vom spune că evenimentele A, B sînt *incompatibile*. Cu alte cuvinte, evenimentele A, B sînt incompatibile dacă acestea nu se pot produce simultan.

Contrarul (complementarul) unui eveniment

Definiţia 7.2.5.

Fie A un eveniment; prin contrarul, (sau complementarul), evenimentului A înţelegem evenimentul care se produce ori de cîte ori nu se produce evenimentul A. Contrarul evenimentului A îl notăm prin $\bar{A}$, (citim „evenimentul non A" sau „contrarul evenimentului A").

Exemplu 7.2.3.

Să notăm cu A evenimentul care se produce atunci cînd, în urma aruncării unui zar, apare faţa marcată cu 1, iar B evenimentul care se produce atunci cînd apare faţa marcată cu 2. Conform definiţiei, $A \cup B$ este evenimentul care se produce atunci cînd, în urma aruncării zarului, apare o faţă marcată cu cel mult două puncte.

Exemplu 7.2.4.

Fie A evenimentul care se produce atunci cînd o persoană sosită într-o staţie de autobuz aşteaptă pînă la sosirea autobuzului necesar cel mult 3 minute, iar B evenimentul care se produce atunci cînd aşteptarea durează cel mult 5 minute. Evenimentul $A \cup B$ se produce cînd persoana aşteaptă în staţie cel mult 5 minute, adică $A \cup B = B$.

Din definiţia reuniunii a două evenimente urmează imediat că următoarele relaţii sînt adevărate:

a) $A \cup A = A$,

b) $A \cup B = B \cup A$,

c) $(A \cup B) \cup C = A \cup (B \cup C)$,

d) $A \cup \Omega = \Omega$, $A \cup \phi = A$

în care A, B, C sînt evenimente oarecare, iar Ω şi Φ sînt, respectiv, evenimentul sigur şi evenimentul imposibil.

Intersecţia a două evenimente

Definiţia 7.2.4.

Fie A, B două evenimente; prin intersecţia lor înţelegem un nou eveniment, care se produce atunci cînd se produce atît evenimentul A cît şi evenimentul B.

Intersecţia evenimentelor A, B o notăm prin $A \cap B$, (citim „evenimentul A intersectat cu B" sau „evenimentul A şi B").

Exemplul 7.2.7.

Considerăm din nou evenimentele A, B din exemplul 7.2.3. Evenimentul $\bar{A}$ este evenimentul care se produce atunci cînd, după aruncarea zarului, apare oarecare dintre feţele marcate cu 2,3,4,5,6, iar evenimentul $\bar{B}$ este evenimentul care se produce atunci cînd după aruncarea zarului apare oricare dintre feţele marcate cu 1,3,4,5,6. Ca şi în cazul reuniunii şi intersecţiei a două evenimente, se constată imediat că următoarele relaţii sînt adevărate:

a) $\bar{\bar{A}} = A$,

b) $\overline{A \cup B} = \bar{A} \cap \bar{B}$, $\overline{A \cap B} = \bar{A} \cup \bar{B}$,

c) $A \cup \bar{A} = \Omega$, $A \cap \bar{A} = \Phi$,

d) $\bar{\Omega} = \Phi$, $\bar{\Phi} = \Omega$

în care A, B sînt evenimente oarecare, iar Ω, Φ, evenimentul sigur, respectiv evenimentul imposibil.

§ 3. CÎMP DE EVENIMENTE

Pînă acum am considerat că un eveniment este pus în legătură cu un anumit fenomen, care poate sau nu să se producă atunci cînd un ansamblu dat de condiţii este îndeplinit. Am mai observat că unei experienţe oarecare, reale sau imaginare, i se poate asocia un ansamblu de evenimente, cu aceste evenimente putîndu-se face unele operaţii (reuniune, intresecţie, contrarul unui eveniment) obţinîndu-se ca rezultat un eveniment din acelaşi ansamblu de evenimente. Pentru a cuprinde ceea ce este semnificativ în problemele în care se pot întîlni evenimente, în sensul acceptat pînă acum, este util ca evenimentele să se definească fără a se apela la diferite situaţii particulare. În cele ce urmează vom defini noţiunile de eveniment şi cîmp de evenimente într-un cadru general.

În cap.I, § 2 am definit corpul de părţi şi am dat unele proprietăţi ale lui.

Un corp de părţi ale mulţimii Ω este o submulţime $K \neq \Phi$ din $\mathcal{P}(\Omega)$ închisă la coplementariere şi reuniune. Reamintim acum cîteva proprietăţi importante ale unui corp de părţi.

Proprietatea 1. Dacă Ω este o mulţime şi K un corp de părţi ale mulţimii Ω atunci $\Phi \in K$, $\Omega \in K$.

Proprietatea 2. Dacă Ω este o mulţime şi K un corp de porţi ale mulţimii Ω, atunci A, $B \in K$ implică $A \cap B \in K$.

Proprietatea 3. Fie Ω o mulţime, K un corp de părţi din Ω şi $A_1, A_2, \ldots, A_n \in K$. Atunci $A_1 \cup \ldots \cup A_n \in K$ şi $A_1 \cap \ldots \cap A_n \in K$ $(\forall) n \geq 2$ $n \in \mathbb{N}$

Observaţia 7.1.

Dacă avem în vedere închiderea la reuniunea unei infinităţi de submulţimi corpul de părţi a fost numit corp borelian de părţi ale mulţimii Ω sau Γ.

Definiţia 7.3.1.

Fie Ω o mulţime oarecare şi K un corp de părţi ale mulţimii Ω. Perechea (Ω, K), formată din mulţimea Ω şi corpul de părţi K, o vom numi cîmp de evenimente.

Dacă (Ω, K) este un cîmp de evenimente, convenim să numim eveniment orice element al mulţimii K, adică prin eveniment al cîmpului de evenimente (Ω, K) se înţelege orice submulţime a mulţimii Ω care se află în corpul de părţi K. După cum am constatat mai înainte, orice corp de părţi ale mulţimii Ω conţine mulţimea vidă, Φ, şi mulţimea Ω; Privite ca elemente ale corpului de părţi K, din cîmpul de evenimente (Ω, K), Φ se va numi eveniment imposibil, iar Ω eveniment sigur.

Această definiţie abstractă pentru cîmpul de evenimente este justificată prin aceea că mulţimea evenimentelor corespunzătoare unei experienţe concrete poate fi pusă în corespondenţă biunivocă cu un corp de părţi ale unei mulţimi convenabil alese, operaţiile dintre evenimente, (reuniunea, intersecţia, contrarul unui eveniment), reflectîndu-se în operaţiile, (reuniune, intersecţie, complementară), dintre elementele corpului de părţi.

Exemplu 7.3.1.

Să presupunem că ne interesează faţa care va apare în urma aruncării

unei monede. Fie A evenimentul care se produce atunci cînd apare „stema", B evenimentul care se produce atunci cînd apare „banul", $A \cup B = \Omega$, (evenimentul sigur), iar $A \cap B = \Phi$, (evenimentul imposibil). Să considerăm acum mulţimile Ω şi K în felul următor:

$\Omega = \{1, 2\}$, $K = \mathcal{P}(\Omega) = \{\{1\}, \{2\}, \Omega, \Phi\}$.

Se constată imediat că (Ω, K) este un cîmp de evenimente (în sensul definiţiei 7.3.1) şi putem stabili corespondenţa între evenimente corespunzătoare experienţei şi K în felul următor:

$A \to \{1\}$, $B \to \{2\}$, $A \cap B = \Phi \to \Phi \in K$, $A \cup B = \Omega \to \{1, 2\} = \Omega$

Remarcăm faptul că mulţimile Ω, K pot fi alese şi în alte moduri, de exemplu: $\Omega = \{a, b, c\}$, $K = \{\{a, b\}, \{c\}, \Omega, \Phi\}$, corespondenţa dintre evenimentele corespunzătoare experienţei date mai sus şi elementele mulţimii Ω luînd-o sub forma:

$A \to \{a, b,\}$, $B \to \{c\}$, $\Phi = A \cap B \to \Phi \in K$, $\Omega = A \cup B \to \{a, b, c\} = \Omega$

Definiţia 7.3.2.

Fie (Ω, K) un cîmp de evenimente. Se numeşte eveniment elementar al cîmpului de evenimente dat orice eveniment A, $A \in K$, $A \neq \Phi$, cu proprietatea că nici-o submulţime propie a mulţimii A nu se află în K, adică din faptul că A, $B \in K$, $B \subset A$, urmează $B = \Phi$ sau $B = A$.

Propunem ca exerciţiu verificarea următoarei caracteristici pentru evenimentele elementare:

Propoziţia 7.3.1.

Fie (Ω, K) un cîmp de evenimente şi $A \in K$, $A \neq \Phi$. Evenimentul A este un eveniment elementar dacă şi numai dacă pentru oricare eveniment B, $B \in K$, are loc $A \cap B = \Phi$ sau $A \cap B = A$.

Definiţia 7.3.3.

Un cîmp de evenimente (Ω, K) se spune că este un *cîmp finit de evenimente* dacă mulţimea K, (corpul de părţi), este o mulţime finită.

Observăm că atunci cînd mulţimea Ω este finită, în mod necesar cîmpul de evenimente (Ω, K) este un cîmp finit de evenimente.

Teorema 7.3.1.

În orice cîmp finit de evenimente există măcar un eveniment elementar.

Demonstraţie: Fie (Ω, K) un cîmp finit de evenimente şi $A \in K$, $A \neq \Phi$. În cazul cînd A este eveniment elementar nu mai avem ce demonstra; dacă A nu este eveniment elementar, atunci există un eveniment A_1, $A_1 \neq \Phi$, A, $A_1 \subset A$. Dacă A_1 este eveniment elementar atunci teorema este demonstrată iar dacă nu, există $A_2 \in K$, $A_2 \neq \Phi$, A_1 astfel ca $A_2 \subset A_1 \subset A$. Deoarece mulţimea K este o mulţime finită este clar că procedînd în continuare în acelaşi mod, după un număr de paşi, vom găsi un eveniment A_p, $A_p \in K$, care este eveniment elementar şi teorema este demonstrată.

Teorema 7.3.2.

Fie (Ω, K) un cîmp finit de evenimente şi $\{A_1, \ldots, A_n\}$ mulţimea tuturor evenimentelor elementare ale acestui cîmp de evenimente. Atunci orice eveniment A, $\Phi \neq A \in K$, se poate scrie ca o reuniune de evenimente elementare.

Demonstraţie: Pentru a demonstra această teoremă, vom arăta mai întîi că reuniunea tuturor evenimentelor elementare dintr-un cîmp finit de evenimente este evenimentul sigur. Să presupunem că $\{A_1, \ldots, A_n\}$ constituie mulţimea tuturor evenimentelor elementare din cîmpul finit de evenimente (Ω, k) şi că reuniunea acestora este diferită de evenimentul sigur Ω, adică $B = \bigcup_{i=1}^{n} A_i \neq \Omega$. Deoarece $A_i \in K$, $i = 1, \ldots, n$, din proprietăţile corpului de părţi rezultă că $B \in K$ şi pe baza definiţiei corpului de părţi urmează $\bar{B} \in K$. Din faptul că $B \neq \Omega$ rezultă $\bar{B} \neq \Phi$ şi atunci, din demonstraţia teoremei 7.3.1 urmează că $\bar{B}$ sau este eveniment elementar sau conţine un eveniment elementar. Deoarece $\bar{B} \cap B = \Phi$, evenimentul elementar conţinut în $\bar{B}$ nu se află în mulţimea $\{A_1, \ldots, A_n\}$, ceea ce contrazice ipoteza că mulţimea $\{A_1, \ldots, A_n\}$ conţine toate evenimentele elementare din cîmpul de evenimente (Ω, K). Prin urmare, presupunerea $\bigcup_{i=1}^{n} A_i \neq \Omega$ nu se poate accepta şi deci $\bigcup_{i=1}^{n} A_i = \Omega$.

Să arătăm acum că orice eveniment, diferit de evenimentul imposibil, dintr-un cîmp finit de evenimente, se poate scrie ca o reuniune de evenimente elementare. Fie (Ω, K) un cîmp finit de evenimente, $\{A_1, \ldots, A_N\}$ mulţimea tuturor evenimentelor elementare din acest cîmp de evenimente şi $A \in K$. Deoarece $\bigcup_{i=1}^{n} A_i = \Omega$ putem scrie succesiv:

$$A = A \cap (A_1 \cup \ldots \cup A_n)$$

$$A = (A \cap A_1) \cup (A \cap A_2) \cup \ldots \cup (A \cap A_n)$$

Ţinînd cont de faptul că A_i, $i = 1, \ldots, n$ sînt evenimente elementare, urmează că pentru orice i, $i \in \{1, \ldots, n\}$, sau $A \cap A_i = A_i$ sau $A \cap A_i = \Phi$ (vezi propoziţia 3.3.1). Pe de altă parte, există i, $i \in \{1, \ldots, n\}$, astfel ca $A \cap A_i \neq \Phi$, pentru că în caz contrar $A = \Phi$, ceea ce a fost exclus de la început. Pe baza celor de mai sus rezultă că există o mulţime $\{i_1, i_2, \ldots, i_p\} \subset \{1, \ldots, n\}$ astfel încît să aibă loc relaţia:

$$A = A_{i_1} \cup A_{i_2} \cup \ldots \cup A_{i_p}$$

şi teorema este demonstrată.

Să observăm că teorema 7.3.2 arată că un cîmp finit de evenimente este în întregime caracterizat de mulţimea evenimentelor elementare ale acestui cîmp.

Definiţia 7.3.3.'

Prin cîmp borelian de evenimente se înţelege un cîmp de evenimente (Ω, K) în care corpul de părţi este un corp borelian de părţi ale mulţimii Ω.

Observaţia 7.2. Din definiţia 7.3.3. şi faptul că într-un cîmp finit de evenimente, (Ω, K), corpul de părţi K este finit rezultă imediat că un cîmp finit de evenimente este în acelaşi timp cîmp borelian de evenimente.

În încheiere mai menţionăm faptul că atunci cînd A, B sînt două evenimente dintr-un cîmp de evenimente (Ω, K) astfel încît $A \cap B = \Phi$ se spune că evenimentele A, B, sînt *incompatibile*.

Exerciţii:

1. Se aruncă o monedă de două ori consecutiv şi ne interesează care dintre feţele monedei va apare după fiecare aruncare. Construiţi un cîmp de evenimente, (Ω, K), corespunzător acestei experienţe şi puneţi în evidenţă evenimentele elementare ale acestui cîmp de evenimente.

2. Se aruncă un zar de două ori consecutiv şi ne interesează faţa care apare după fiecare aruncare. Descrieţi evenimentele elementare corespunzătoare acestei experienţe. Caracterizaţi cîmpul de evenimente, asociat experienţei, prin evenimentele sale elementare.

3. Arătaţi că două evenimente A, B, $A \neq B$, dintr-un cîmp de evenimente (Ω, K) sînt incompatibile.

4. Să presupunem că (Ω, K) este un cîmp de evenimente şi A_i, $i = 1, \ldots, n$, sînt evenimente din acest cîmp de evenimente satisfăcînd condiţiile:

a) $A_i \neq \Phi$, $i = 1, \ldots, n$,

b) $A_i \cap A_j = \Phi$, $i \neq j$,

c) $\bigcup_{i=1}^{n} A_i = \Omega$.

Se poate spune că evenimentele A_i, $i = 1, \ldots, n$ sînt evenimentele elementare? Justificaţi concluzia şi printr-un exemplu.

5. Într-un grup de 4 persoane două au vîrsta de 20 ani iar celelalte două au vîrsta de 21 ani. Din acest grup aleg la întîmplare două persoane. În ipoteza că ne interesează ce vîrstă are fiecare din cele două persoane care urmează a fi alese, descrieţi evenimentele elementare corespunzătoare acestei experienţe şi caracterizaţi cîmpul de evenimente asociat experienţei prin evenimentele sale elementare.

6. Fie (Ω, K) un cîmp de evenimente şi B, $B \neq \Phi$, un eveniment din acest cîmp de evenimente. Dacă notăm $K_B = \{A \cap B | A \in K\}$, arătaţi că perechea (B, K_B) este deasemenea un cîmp de evenimente.

§ 4. PROBABILITATE. CÎMP DE PROBABILITATE

După noţiunile de evenimente şi cîmp de evenimente, următoarea noţiune fundamentală pentru calculul probabilităţilor este noţiunea de probabilitate. După cum am văzut, un eveniment poate fi pus în legătură cu producerea unui anumit fenomen atunci cînd este îndeplinit un ansamblu dat de condiţii. Din punct de vedere practic, a spune că un astfel de eveniment se poate sau nu produce, nu constituie un răspuns satisfăcător, precizarea unei măsuri în care condiţiile date asigură producerea fenomenului respectiv fiind necesară. Pentru ilustrare, să presupunem că durata de utilizare a unei anumite piese de schimb variază între 50 şi 500 eveniment elementar al acestui cîmp de evenimente se mai numeşte şi rezultatat posibil al experienţei sau *caz posibil*. Vom spune că rezultatele, (cazurile), posibile ale experienţei sînt egal posibile dacă nu există motive obiective care să ne permită să spunem că unele din aceste rezultate posibile ar putea avea şanse de apariţie mai mari decît altele, în urma efectuării experinţei respective.

În această situaţie, evenimentele elementare corespunzătoare se vor numi evenimente egal posibile, sau *cazuri egal posibile*.

După cum am văzut în paragraful precedent, orice eveniment dintr-un cîmp finit de evenimente se poate scrie ca o reuniune de evenimente elementare. Dacă A este un eveniment dintr-un astfel de cîmp de evenimente şi $A = A_1 \cup \ldots \cup A_p$, unde A_i, $i = 1, \ldots p$, sînt evenimente elementare, atunci evenimentele A_i, $i = 1, \ldots, p$, se mai numesc şi evenimente elementare favorabile producerii evenimentului A, sau *cazuri favorabile* producerii evenimentului A.

Exemplul 7.4.4.

Să presupunem că se aruncă un zar obişnuit, marcat pe fiecare dintre feţe cu puncte de la 1 la 6 şi ne interesează faţa care va apare în urma aruncării zarului. Să notăm cu A evenimentul care se produce cînd apare o faţă marcată cu cel mult 3 puncte. Un eveniment elementar asociat acestei experienţe corespunde apariţiei uneia dintre cele 6 feţe ale zarului; să notăm cu A_i, $1 \le i \le 6$, evenimentul care se produce atunci cînd, în urma aruncării zarului, apare faţa marcată cu i puncte, $1 \le i \le 6$. Evenimentele A_i, $i = 1, \ldots, 6$, sînt evenimente egal posibile dacă zarul este perfect echilibrat,

astfel încît oricare dintre cele 6 feţe ale zarului să aibă aceleaşi şanse de a apare (în acest caz numărul cazurilor egal posibile ar fi egal cu 6).

Pe de altă parte, este clar că evenimentul A, considerat mai înainte, se poate scrie sub forma: $A = A_1 \cup A_2 \cup A_3$ şi atunci fiecare dintre evenimentele A_1, A_2, A_3 constituie un caz favorabil producerii evenimentului A (numărul cazurilor favorabile producerii evenimentului A este egal cu 3). Vom da acum definiţia, numită clasică, a probabilităţii.

Definiţia 7.4.1.

Fie A un eveniment dintr-un cîmp finit de evenimente, asociat unei experienţe, în care evenimentele elementare sînt egal posibile. Prin definiţie, probabilitatea evenimentului A, notată $P(A)$, se consideră ca fiind raportul dintre numărul cazurilor (rezultatelor) favorabile producerii evenimentului A şi numărul cazurilor (rezultatelor) egal posibile, adică

$$P(A) = \frac{k}{n}$$

în care n este numărul cazurilor egal posibile (numărul tuturor evenimentelor elementare din cîmpul de evenimente considerat), iar k este numărul cazurilor favorabile producerii evenimentului A, (numărul evenimentelor elementare a căror reuniune constituie evenimentul A).

Înainte de a da exemple este bine să subliniem faptul că această definiţie a probabilităţii unui eveniment poate fi folosită numai atunci cînd sînt îndeplinite următoarele două condiţii: i) cîmpul de evenimente este finit şi ii) evenimentele elementare ale acestui cîmp sînt egal posibile, adică oricare astfel de evenimente au aceleaşi şanse de a se produce.

Exemplu 7.4.5.

Să considerăm o urnă U în care se află 3 bile albe şi 5 bile negre şi fie A evenimentul care se produce atunci cînd, luînd la întîmplare o bilă din această urnă, constatăm că aceasta este albă. Deoarece se alege o bilă din cele $3+5=8$ bile existente în urnă, dacă fiecare dintre acestea are aceeaşi şansă de a fi aleasă, numărul cazurilor egal posibile este 8 şi deoarece pentru a se produce evenimentul A trebuie aleasă una din cele 3 bile albe existente în urnă, numărul cazurilor favorabile producerii evenimentului A este 3. Prin

urmare, conform definiţiei clasice a probabilităţii, $P(A) = \frac{3}{8}$.

Să observăm că în cadrul acestei experienţe fiecare eveniment elementar corespunde extragerii unei anumite bile din cele 8 bile existente în urnă, aceste evenimente elementare putînd fi considerate ca fiind egal posibile. Acestei experienţe am putea să-i asociem şi un cîmp de evenimente cu numai două evenimente elementare: A evenimentul care se produce atunci cînd bila extrasă este albă şi B evenimentul care se produce atunci cînd bila extrasă este neagră, dar în acestă situaţie cele două cazuri posibile (bila extrasă să fie albă sau bila extrasă să fie neagră) nu mai sînt egal posibile, fiind destul de clar că sînt mai mari şanse de a extrage o bilă neagră decît de a extrage o bilă albă.

Exemplul 7.4.6.

Să presupunem că vrem să determinăm probabilitatea ca, din două aruncări consecutive ale unui zar, suma punctelor obţinute să fie egală cu 3.

Deoarece zarul are 6 feţe, în fiecare aruncare putînd apare oricare dintre cele 6 feţe, dacă oricare faţă are aceleaşi şanse de a apare, este clar că numărul cazurilor egal posibile este $6\times6=36$. Dacă notăm cu A evenimentul care se produce atunci cînd suma punctelor obţinute în cele două aruncări este 3 atunci numărul cazurilor favorabile producerii evenimentului A va fi egal cu numărul acelor cazuri, din cele 36 posibile, în care suma dintre primul şi al doilea este egal cu 3: acestea vor fi cazurile în care se obţine 1 în prima aruncare şi 2 în cea de a doua aruncare, sau 2 în prima aruncare şi 1 în cea de a doua aruncare, deci în total două cazuri favorabile producerii evenimentului A. Prin urmare $P(A) = \frac{2}{36}$, este echivalent $P(A) = \frac{1}{18}$.

Avînd în vedere modul în care a fost definită mai sus probabilitatea unui eveniment, dintr-un cîmp finit de evenimente, putem dovedi următoarea teoremă:

Teorema 7.4.1.

Fie (Ω, K) un cîmp finit de evenimente în care evenimentele elementare, $A_1, A_2, \ldots, A_n$, sînt considerate egal posibile. Atunci, dacă prin probabilitatea unui eveniment oarecare din acest cîmp de evenimente se

înţelege raportul dintre numărul cazurilor favorabile producerii evenimetului respectiv şi numărul cazurilor egal posibie, sînt adevărate următoarele afirmaţii:

a) Oricare ar fi un eveniment A, are loc $P(A) \geq 0$

b) $P(\Omega) = 1$, unde prin Ω se înţelege, ca de obicei, evenimentul sigur.

c) Dacă A şi B sînt două evenimente incompatibile din cîmpul de evenimente considerat, atunci $P(A \cup B) = P(A) + P(B)$.

Demonstraţie: Prima dintre afirmaţii rezultă din faptul că raportul dintre numărul nenegativ (numărul cazurilor favorabile producerii evenimetului A) şi un număr pozitiv (numărul cazurilor egal posibile) este întotdeauna nenegativ. Cea de a doua dintre afirmaţii este o consecinţă imediată a faptului că numărul cazurilor favorabile producerii evenimentului sigur este egal cu numărul cazurilor egal posibile (a se vedea demonstraţia teoremei 7.3.2).

În sfîrşit, pentru a justifica a treia afirmaţie să presupunem că $A = A_{i_1} \cup \ldots \cup A_{i_k}$ şi $B = A_{j_1} \cup \ldots \cup A_{j_r}$, unde $A_{i_1}, \ldots, A_{i_k}$, $A_{j_1}, \ldots, A_{j_r}$ sînt evenimente elementare din cîmpul de evenimente considerat (vezi teorema 7.3.2). Conform definiţiei clasice a probabilităţii, putem scrie $P(A) = \frac{k}{n}$, $P(B) = \frac{r}{n}$.

Pentru a găsi probabilitatea evenimentului $A \cup B$ este nevoie să determinăm numărul cazurilor favorabile producerii acestui eveniment, numărul cazurilor egal posibil fiind deasemeni n. Dar, din cele de mai sus, putem scrie:

$$A \cup B = (A_{i_1} \cup \ldots \cup A_{i_k}) \cup (A_{j_1} \cup \ldots \cup A_{j_r})$$

şi deci numărul cazurilor favorabile producerii evenimentului $A \cup B$ este cel mult $(k + r)$. Cum evenimentele A, B au fost presupuse incompatibile, $A \cap B = \Phi$, urmează că nu există cazuri favorabile atît producerii evenimentului A cît şi a evenimentului B şi, prin urmare, numărul cazurilor favorabile producerii evenimentului $A \cup B$ este exact $(k + r)$. Deci:

$$P(A \cup B) = \frac{k+r}{n} = \frac{k}{n} + \frac{r}{n}$$

şi cum $P(A) = \frac{k}{n}$, $P(B) = \frac{r}{n}$ urmează:

$$P(A \cup B) = P(A) + P(B)$$

ceea ce demonstrează complet teorema.

Probabilităţi condiţionate (abordare intuitivă)

Există situaţii în care sîntem interesaţi să ştim în ce măsură anumite evenimente sînt în interdependenţă cu alte evenimete. De cele mai multe ori, în astfel de situaţii se impune să determinăm probabilitatea producerii unui eveniment presupunînd că un alt eveniment se produce. În exemplele următoare sînt ilustrate astfel de situaţii:

Exemplul 7.4.7.

Să presupunem că în două urne, U_1, U_2, se află bile albe şi bile negre, fiecare dintre urne conţinînd cîte o bilă albă şi una neagră şi să considerăm că facem o experienţă în felul următor: luăm la întîmplare o bilă din urna U_1 şi o introducem în urna U_2, după care luăm la întîmplare o bilă din urna U_2.

Fie A_1, evenimentul care se produce atunci cînd bila extrasă din urna U_1 este albă,iar A_2 evenimentul care se produce atunci cînd bila extrasă din urna U_2 este albă. Este clar că şansele de producere a evenimentului A_2 în ipoteza că evenimentul A_1 se produce, sînt mai mari decît în cazul cînd evenimentul A_1 nu se produce, deoarece în primul caz în urna U_2 creşte numărul de bile albe, iar în cel de al doilea caz creşte numărul de bile negre. Considerînd definiţia clasică a probabilităţii, probabilitatea producerii evenimentului A_2, în ipoteza că evenimentul A_1 se produce va fi egală cu 2/3, deoarece în urna U_2 se vor afla 3 bile din care două albe, iar probabilitatea producerii evenimentului A_2, în ipoteza că evenimentul A_1 nu se produce, va

fi egală cu 1/3, deoarece, atunci cînd nu se produce evenimentul A_1 din urna U_1 se ia o bilă neagră care introdusă în urna U_2 face ca urna U_2 să conţină o bilă albă şi două bile negre înaintea extragerii unei bile din U_2.

În general, presupunînd că A şi B, $B \neq \Phi$, sînt două evenimente dintr-un cîmp de evenimente corespunzător unei experienţe concrete, vom nota prin A/B evenimentul care se produce atunci cînd, în ipoteza că se produce evenimentul B, se produce şi evenimentul A. Evenimentul A/B se numeşte „evenimentul A condiţionat de B". Remarcăm faptul că în evenimentul A/B nu este implicată o ordine a producerii evenimentelor A şi B. Revenind acum la cazul general, dacă A, B sînt două evenimente atunci probabilitatea evenimentului A/B, notată $P(A/B)$, se numeşte probabilitate a evenimentului A condiţionată de evenimentul B. Din observaţiile făcute mai sus, relativ la evenimentele A_1/A_2 descrise în exemplul 7.4.7, urmează că $P(A_2/A_1) = \frac{2}{3}$ şi $P(A_2/\bar{A}_1) = \frac{1}{3}$, unde $\bar{A}_1$ este contrarul evenimentului A_1.

Să vedem acum putem găsi probabilitatea $P(A/B)$ atunci cînd A, B sînt două evenimente dintr-un cîmp finit de evenimente, în care probabilitatea unui eveniment este considerată conform definiţiei clasice a probabilităţii.

Dacă A, B sînt evenimente dintr-un cîmp finit de evenimente în care se află nt evenimente elementare egal posibile şi dacă $p_1, p_1 > 0$, este numărul cazurilor favorabile producerii evenimentului B atunci $P(B) = \frac{p_1}{n} > 0$.

Presupunînd acum că evenimentul B se produce, cîmpul de evenimente iniţial se transformă într-un cîmp de evenimente în care evenimentele elementare sînt numai acele evenimente elementare, din cîmpul de evenimente iniţial, favorabile producerii evenimentului B. Prin urmare, deoarece evenimentul A/B se produce atunci cînd se produce A cu condiţia producerii evenimentului B, putem spune că probabilitatea A/B este aceeaşi cu probabilitatea $A \cap B$, $A \cap B$ fiind considerat ca eveniment în acest nou cîmp de evenimente. Fie acum p_2 numărul cazurilor favorabile producerii evenimentului $A \cap B$ în cîmpul de evenimente iniţial; atunci putem scrie

$P(A/B) = \frac{p_2}{p_1}$. Dar $P(B) = \frac{p_1}{n}$ iar $P(A \cap B) = \frac{p_2}{n}$, de unde urmează:

$$P(A/B) = \frac{P(A \cap B)}{P(B)}$$

Vom ilustra raţionamentul de mai sus folosind datele exemplului 7.4.7.

Să notăm cu a_1 bila albă şi cu n_1 bila neagră din urna U_1, iar cu a_2, n_2 bila albă, respectiv bila neagră din urna U_2. Ţinînd cont de modul în care se extrage bilele din cele două urne rezultă că toate situaţiile posibile sînt descrise de perechile:

$$(a_1, a_1),\ (a_1, a_2),\ (a_1, n_2),\ (n_1, n_1),\ (n_1, a_2),\ (n_1, n_2),$$

în care primul element al unei perechi reprezintă bilă extrasă din urna U_1 iar cel de al doilea element reprezintă bila extrasă din urna U_2. Aceste perechi corespund evenimentelor elementare de evenimente asociate experienţei. Presupunînd că s-ar produce evenimentul A_1, (extragerea unei bile albe din urna U_1), atunci dintre toate cele şase evenimente elementare iniţiale rămîn posibile 3 evenimente elementare şi anume cele corespunzătoare perechilor (a_1, a_1), (a_1, a_2), (a_1, n_2), care reprezintă cazurile favorabile producerii evenimentului A_1. Dintre aceste trei cazuri numai două, (a_1, a_1), (a_1, a_2), sînt favorabile şi producerii evenimentului A_2, acestea fiind aceleaşi cu cazurile favorabile producerii evenimentului $A_1 \cap A_2$. Deci $\mathcal{P}(A_2/A_1) = \frac{2}{3}$, fapt care a fost constatat şi mai înainte. Pentru aceleaşi evenimente, A_1, A_2, să determinăm acum $P(A_1|A_2)$. Din cele stabilite mai sus, putem scrie:

$$P(A_1|A_2) = \frac{P(A_1 \cap A_2)}{P(A_2)}$$

şi deoarece $P(A_1 \cap A_2) = \frac{2}{6}$, $P(A_2) = \frac{3}{6}$ rezultă că $P(A_1|A_2) = \frac{2}{3}$.

Observaţia 7.3.: Dacă A este un eveniment dintr-un cîmp de evenimente asociat unei experienţe concrete, atunci prin probalitatea evenimentului A, $P(A)$, înţelegem o măsură *a priori* a şanselor ca evenimentul A să se producă, altfel spus, $P(A)$ se consideră definită înaintea efectuării propriu-zise a experienţei. În cazul cînd A, B sînt două evenimente din cîmpul de evenimente asociat experienţei date, atunci $P(A/B)$ este o *măsură a priori* a şanselor ca evenimnentul A să se producă în sensul că, dacă după efectuarea experienţei s-ar constata că evenimentul B s-a produs, $P(A/B)$ reprezintă o măsură a şanselor ca şi evenimentul A să se fi produs (fără a se fi putut constata dacă evenimentul A s-a produs sau nu).

Exerciţii:

1. O monedă este aruncată de două ori consecutiv. Care este probabilitatea ca: a) numai în una din aruncări să apară „marca", b) în cel puţin una din aruncări să apară „marca".

2. Se aruncă două două zaruri. Care este probabilitatea ca măcar pe unul din cele două zaruri să apară faţa marcată cu 6 puncte.

3. Într-o urnă se află două bile negre şi o bilă albă. Din această urnă se extrag cuccesiv două bile, prima bilă extrasă fiind reintrodusă în urnă înainte de a extrage cea de a doua bilă. Să se determine ca cel puţin una din cele două bile extrase să aibă culoarea neagră şi probabilitatea ca cel puţin una dintre cele două extrase să aibă culoarea albă.

4. În condiţiile exerciţiului precedent, să presupunem că după prima extracţie bila nu se reintroduce în urnă, astfel că după cele două extracţiii în urnă rămîne o singură bilă. Fie A evenimentul care se produce atunci cînd prima bilă extrasă este de culoare albă şi B evenimentul care se produce atunci cînd cea de a doua bilă extrasă este de culoare albă. să se determine probabilitatea evenimentelor A, B, $A \cup B$, $A \cap B$, $\bar{A} \cap B$, $A \cap \bar{B}$, $\bar{A} \cap \bar{B}$ şi probabilităţile condiţionate $P(A|B)$, $P(B|A)$, $P(\bar{A}|A)$, $P(A|\bar{B})$.

5. O ţintă în formă de disc este împărţite în trei zone prin două cercuri concentrice avînd centrul în centrul discului. Dacă se trage o lovitură asupra acestei ţinte, ne interesează zona în care lovitura atinge ţinta. Descrieţi evenimentele elementare corespunzăre acestei experienţe. Se poate accepta că aceste evenimente elementare sînt egal posibile?

Definiţia axiomatică a probabilităţii. Cîmp de probabilitate

După cum am văzut în paragraful precedent, definiţia clasică a probabilităţii nu poate fi utilizată decît în cazul cîmpurilor finite de evenimente în care evenimentele sînt egal posibile. Există însă probleme concrete în care utilizarea unui cîmp finit de evenimente se poate dovedi neadecvată sau, chiar dacă se utilizează un cîmp finit de evenimente, evenimentele elementare nu sînt egal posibile. Din acest motiv este necesar să se definească probabilitatea unui eveniment în cazul cînd acesta aparţine unui cîmp de evenimente oarecare. Ţinînd cont de condiţiile naturale pe care ar trebui să le îndeplinească măsura realizării unui eveniment, sîntem conduşi la următoarea definiţie a probabilităţii unui eveniment dintr-un cîmp de evenimente dat, numită definiţia axiomatică a probabilităţii:

Definiţia 7.4.2.

Fie (Ω, K) un cîmp de evenimente; prin probabilitate peste acest cîmp de evenimente înţelegem o funcţie $P: K \to \mathbb{R}$, care îndeplineşte condiţiile:

1. Pentru oricare A, $A \in K$, are loc $P(A) \geq 0$.
2. $P(\Omega) = 1$
3. Dacă $A, B \in K$, $A \cap B = \Phi$, atunci $P(A \cup B) = P(A) + P(B)$.

Dacă (Ω, k) este un cîmp de evenimente, iar $P: K \to \mathbb{R}$ este o probabilitate definită peste acest cîmp de evenimente atunci vom spune că $P(A), A \in K$, este probabilitatea evenimentului A.

Să observăm că probabilitatea definită conform definiţiei clasice îndeplineşte condiţiile din această definiţie, (vezi T 7.4.1) şi, prin urmare, toate proprietăţile probabilităţii definită axiomatic vor fi valabile şi pentru probabilitatea definită clasic. Mai notăm faptul că peste un cîmp de evenimente se pot defini mai multe probabilităţi, în sensul definiţiei axiomatice a probabilităţii, fără ca aceasta să vină în contradicţie cu realitatea, ci dimpotrivă. Să exemplificăm: presupunem că aruncăm o monedă şi ne interesează probabilitatea ca în urma aruncării acestei monede să apară „stema". Conform definiţiei clasice probabilitatea acestui eveniment este 1/2, însă în cazul cînd moneda este *perfect echilibrată*, altfel neavînd posibilitatea să precizăm probabilitatea acestui eveniment, aceasta putînd fi, în funcţie de proprietăţile monedei şi

condiţiile de aruncare, un număr rel cuprins între 0 şi 1. Dacă se consideră cîmpul de evenimente (Ω, K), unde $\Omega = \{1,2\}$, $K = \{\{1\},\{2\},\Phi,\Omega\}$, asociat acestei experienţe, iar evenimentul corespunzător apariţiei „stemei" îi asociem $\{1\} \in K$ şi apariţia „banului" $\{2\} \in K$ atunci definiţia axiomatică a probabilităţii ne permite să definim $P: K \to \mathbb{R}$ astfel încît $P(\{1\}) = p$, $P(\{2\}) = 1 - p$, $P(\Omega) = 1$, $P(\Phi) = 0$, p fiind un număr real cuprins între 0 şi 1, număr care, eventual, ar urma să fie precizat.

Avînd definite noţiunile de cîmp de evenimente şi de probabilitate peste un cîmp de evenimente, putem defini acum noţiunea de *cîmp de probabilitate.*

Definiţia 7.4.3.

Se numeşte cîmp de probabilitate un triplet de forma (Ω, K, P), în care (Ω, K) este un cîmp de evenimente, iar P este o probabilitate definită peste (Ω, K). Vom dovedi în continuare că, în cazul cînd (Ω, K, P) este un cîmp de propabilitate, probabilitatea P are următoarele proprietăţi simple:

1. Pentru orice eveniment A, $A \in K$, are loc $P(\bar{A}) = 1 - P(A)$.
2. $P(\Phi) = 0$.
3. Pentru orice eveniment A, $A \in K$, are loc $0 \le P(A) \le 1$.
4. Dacă $A, B \in K$, $A \subset B$, atunci $P(A) \le P(B)$.
5. Dacă A, $B \in K$ atunci $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.
6. Dacă evenimentele $A_1, \ldots, A_n \in K$ sînt incompatibile două cîte două atunci:

$$P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i)$$

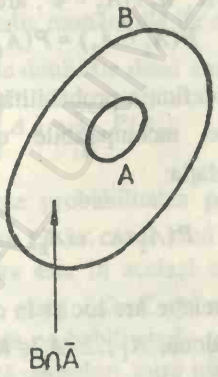
Demonstraţie:

1. Fie $A \in K$; atunci $\bar{A} \in K$ şi $A \cup \bar{A} = \Omega$; deci $P(A \cup \bar{A}) = P(\Omega)$. Conform condiţiei 2 din defniţia axiomatică a probabilităţii $P(\Omega) = 1$, iar din condiţia 3 urmează $P(A \cup \bar{A}) = P(A) + P(\bar{A})$, deoarece $A \cap \bar{A} = \Phi$. Prin urmare $P(A) + P(\bar{A}) = 1$, de unde $P(\bar{A}) = 1 - P(A)$.

2. Deoarece $\Phi = \bar{\Omega}$, putem scrie $P(\Phi) = P(\bar{\Omega})$ şi ţinînd cont de proprietatea 1 şi axioma 2 din definiţia probabilităţii, urmează $P(\Phi) = 1 - P(\Omega) = 1 - 1 = 0$.

3. Faptul că $P(A) \geq 0$, $(\forall) A \in K$, rezultă din prima axiomă a definiţiei probabilităţii. Din prima proprietate, demonstrată mai sus, rezultă $P(A) = 1 - P(\bar{A})$, $(\forall) A \in K$, şi cum $P(\bar{A}) \geq 0$ urmează $P(A) \leq 1$, $(\forall) A \in K$, ceea ce completează demonstraţia proprietăţii 3.

4. Fie $A, B \in K$, $A \subset B$; atunci putem scrie $B = A \cup (B \cap \bar{A})$ şi $P(B) = P(A \cup (B \cap \bar{A}))$. Deoarece $A \cap (B \cap \bar{A}) = \Phi$, pe baza axiomei 3 din definiţia probabilităţii, urmează $P(A \cup (B \cap \bar{A})) = P(A) + P(B \cap \bar{A})$ şi deci putem scrie $P(B) = P(A) + P(B \cap \bar{A})$. Dar $P(B \cap \bar{A}) \geq 0$ şi atunci este clar că $\mathcal{P}(B) \geq \mathcal{P}(A)$, ceea ce demonstrează proprietatea 4. Raţionamentul este mai uşor de urmărit pe figura:

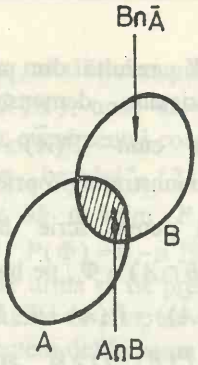


5. Să considerăm $A, B \in K$. Ţinînd cont de proprietăţile operaţiilor cu mulţimi, putem scrie $A \cup B = A \cup (B \cap \bar{A})$, $B = (A \cap B) \cup (B \cap \bar{A})$ şi deoarece $A \cap (B \cap \bar{A}) = \Phi$, $(A \cap B) \cap (B \cap \bar{A}) = \Phi$, pe baza axiomei 3 din definiţia probabilităţii, urmează

$$P(A \cup B) = P(A) + P(B \cap \bar{A}),\ P(B) = P(A \cap B) + P(B \cap \bar{A})$$

Prin urmare $P(A \cup B) - P(B) = P(A) - P(A \cap B)$, care este echivalentă cu relaţia enunţată în proprietatea 5 şi deci proprietatea 5 este demonstrată.

Ilustrăm cu figura de pe pagina următoare:



6. Demonstrația acestei proprietăți o vom face prin inducție după numărul n al evenimentelor considerate. În cazul cînd $n = 2$, pentru orice evenimente incompatibile, $A_1, A_2 \in K$, $A_1 \cap A_2 = \Phi$, are loc relația:

$$P(A_1 \cup A_2) = P(A_1) + P(A_2)$$

pe baza axiomei 3 din definiția probabilității. Să presupunem acum că pentru oricare k evenimente incompatibile două cîte două, $A_1, \ldots, A_k \in K$, $2 \le k \le n-1$, are loc relația:

$$P(A_1 \cup \ldots \cup A_k) = \sum_{i=1}^{k} P(A_i)$$

și să vedem că această relație are loc și în cazul cînd $k = n$. Fie n evenimente incompatibile două cîte două, $A_1, \ldots, A_n \in K$.

Atunci:

$$\bigcup_{i=1}^{n} A_i = A_1 \cup \ldots \cup A_{n-1} = (A_1 \cup \ldots \cup A_{n-1}) \cup A_n$$

și deci putem scrie:

$$P(\bigcup_{i=1}^{n} A_i) = P((A_1 \cup \ldots \cup A_{n-1}) \cup A_n) = P(A_1 \cup \ldots \cup A_{n-1}) + P(A_n)$$

deoarece, din faptul că evenimentele $A_1, \ldots, A_n$ sînt incompatibile două cîte două, evenimentele $A_1 \cup \ldots \cup A_{n-1}$ și A_n sînt incompatibile. Din ipoteza inductivă urmează că:

$$P(A_1 \cup \ldots \cup A_{n-1}) = P(A_1) + \ldots + P(A_{n-1})$$

care împreună cu relaţia precedentă justifică proprietatea. Pentru o analiză mai profundă a unor probleme cu caracter teoretic sau/şi practic este necesar ca pe lîngă noţiunea de cîmp de probabilitate definită mai sus să se considere noţiunea de *cîmp boreliean de probabilitate* a cărei definiţie o prezentăm în continuare.

Definiţia 7.4.4.

Se numeşte cîmp borelian de probabilitate un triplet de forma (Ω, K, P), în care (Ω, K) este un cîmp borelian de evenimente iar $\mathcal{P}$ este o funcţie, $P: K \to \mathbb{R}$, ce îndeplineşte condiţiile:

1'. $P(A) \geq 0$, $(\forall) A \in K$,

2'. $P(\Omega) = 1$

3'. Dacă $(A_i)_{i \in I}$ este o familie numărabilă de evenimente din cîmpul de evenimente (Ω, K), incompatibile două cîte două atunci:

$$P(\bigcup_{i \in I} A_i) = \sum_{i \in I} P(A_i)$$

În acest caz P se numeşte probabilitatea peste cîmpul borelian de evenimente (Ω, K). Să observăm că în cazul cînd (Ω, K, P) este un *cîmp finit* de probabilitate, atunci acesta este în acelaşi timp şi cîmp borelian de probabilitate, iar pe de altă parte, dacă (Ω, K, P) este un cîmp borelian de probabilitate, atunci este şi cîmp de probabilitate în sensul definiţiei 7.4.3. Cu alte cuvinte, putem spune că există deosebiri între noţiunea de cîmp de probabilitate şi noţiunea de cîmp borelian de probabilitate numai în cazul cînd cîmpul de evenimente corespunzător nu este un cîmp finit de evenimente, orice proprietate a unui cîmp de probabilitate fiind adevărată şi într-un cîmp borelian de probabilitate. În particular, proprietăţile 1-6 ale probabilităţii $\mathcal{P}$ dintr-un cîmp de probabilitate sînt adevărate şi pentru probabilitatea $\mathcal{P}$ dintr-un cîmp borelian de probabilitate, (proprietatea 6 este chiar mai slabă decît condiţia 3' din definiţia 7.4.4). Vom conveni ca de acum înainte prin cîmp de probabilitate să înţelegem cîmp borelian de proprietate, cu excepţia cazurilor cînd se precizează altfel.

Probabilităţi condiţionate

În paragraful precedent, după ce s-a dat definiţia clasică a probabilităţii unui eveniment am văzut cum se poate defini probabilitatea producerii unui eveniment, A, condiţionat de producerea unui alt eveniment, B. Exemplul 7.4.7. şi analiza ulterioară a acestuia a pus în evidenţă modul în care se poate interpreta o astfel de probabilitate condiţionată în cazul unor evenimente concrete. Prin analogie, vom defini acum probabilitatea în cazul unui cîmp de probabilitate oarecare (Ω, K, P).

Definiţia 7.4.5.

Fie (Ω, K, P) un cîmp de probabilitate şi $B \in K$ astfel ca $P(B) \neq 0$. Dacă A este un eveniment oarecare din K, atunci prin ~~proprietatea~~ probabilitatea evenimentului A condiţionată de evenimentul B, notată $P(A|B)$, se înţelege raportul dintre probabilitatea intersecţiei evenimentelor A, B şi probabilitatea evenimentului B, adică:

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

În cele ce urmează vom justifica unele proprietăţi ale probabilităţii condiţionate definite mai sus. Remarcăm faptul că relaţia de definiţie a probabilităţilor condiţionate din defifiniţia 7.4.5. este identică cu proprietatea, evidenţiată anterior, pentru probabilitatea condiţionată în accepţiune intuitivă.

Avînd în vedere şi observaţia 7.3, vom considera întotdeauna că, pentru un cîmp de probabilitate oarecare (Ω, K, P), probabilitatea condiţionată $P(A|B)$ reprezintă probabilitatea producerii evenimentului A în ipoteza că s-ar produce evenimentul B. Deasemenea, notăm faptul că pentru probabilitatea condiţionată $P(A|B)$ se mai foloseşte notaţia $P_B(A)$.

Dacă (Ω, K, P) este un cîmp de probabilitate, iar B un eveniment din acest cîmp de probabilitate astfel încît $P(B) \neq 0$ atunci:

1. $P(\Omega|B) = 1$, $P(B|B) = 1$
2. Dacă $A_1, A_2 \in K$, $A_1 \cap A_2 = \Phi$ atunci:

$$P(A_1 \cup A_2|B) = P(A_1|B) + P(A_2|B)$$

3. Dacă $A \in K$, $P(A) \neq 0$, atunci:

$$P(A)\,P(B|A) = P(B)\,P(A|B) = P(A \cap B)$$

Demonstraţie:

1. Deoarece $\Omega \cap B = B$ şi $B \cap B = B$, rezultă imediat din definiţia probabilităţii condiţionate că $P(\Omega|B) = 1$, $P(B|B) = 1$.

2. Fie $A_1, A_2 \in K$, $A_1 \cap A_2 = \Phi$. Conform definiţiei probabilităţii condiţionate putem scrie:

$$P(A_1 \cup A_2|B) = \frac{P((A_1 \cup A_2) \cap B)}{P(B)}$$

Deoarece $(A_1 \cup A_2) \cap B = (A_1 \cap B) \cup (A_2 \cap B)$, iar incompatibilitatea evenimentelor A_1, A_2 atrage incompatibilitatea evenimentelor $A_1 \cap B$ şi $A_2 \cap B$ urmează:

$$P(A_1 \cup A_2|B) = \frac{P(A_1 \cap B) + P(A_2 \cap B)}{P(B)}$$

ceea ce este echivalent cu:

$$P(A_1 \cup A_2|B) = P(A_1|B) + P(A_2|B)$$

3. Presupunînd evenimentul $A \in K$ astfel încît $P(A) \neq 0$ atunci putem defini probabilitatea condiţionată:

$$P(B|A) = \frac{P(A \cap B)}{P(A)}$$

şi deoarece:

$$\mathcal{P}(A|B) = \frac{\mathcal{P}(A \cap B)}{\mathcal{P}(B)}$$

rezultă că $P(A)\,P(B|A) = P(A \cap B)$ şi totodată $P(B)\,P(A|B) = P(A \cap B)$ ceea ce implică $P(A)\,P(B|A) = P(B)\,P(A|B)$ şi cu aceasta cele trei proprietăţi enunţate mai sus sînt dovedite.

BIBLIOTECA CENTRALĂ UNIVERSITARĂ

Observaţia 7.4: Dacă (Ω, K, P) este un cîmp de probabilitate şi B este un eveniment fixat $B \in K$, aşa fel încît $P(B) \neq 0$, atunci putem considera functia $P_B : K \to \mathbb{R}$ definită prin:

$$P_B(A) = \frac{P(A \cap B)}{P(B)}, \ (\forall) A \in K$$

adică $P_B(A)$ este probabilitatea evenimentului A condiţionată de evenimentul B. Ţinînd cont de proprietăţile 1 şi 2 ale probabilităţilor condiţionate şi de faptul evident că $P_B(A) \geq 0$, $(\forall) A \in K$, rezultă că (Ω, K, P_B) este deasemenea un cîmp de probabilitate. În cazul cînd (Ω, K, P) este un cîmp borelian de probabilitate atunci (Ω, K, P_B) este deasemeni un cîmp borelian de probabilitate. Propunem ca exerciţiu verificarea acestei ultime afirmaţii.

Evenimente independente

Revenind la definiţia intuitivă a noţiunii de eveniment, dacă se consideră două evenimente A, B, este posibil ca producerea unuia dintre acestea să influenţeze într-o anumită măsură producerea celuilalt, sau nu. Dacă evenimentele A şi B sînt astfel încît producerea oricăruia dintre ele nu influenţează în nici un fel producerea celuilalt, atunci în mod natural vom spune că evenimentele A şi B sînt *independente*; în caz contrar vom spune că aceste evenimente sînt *dependente*.

Exemplul 7.4.8.

Să presupunem că într-o urnă se află două bile, una de culoare albă şi una de culoare neagră şi să presupunem că intenţionăm să scoatem din urnă cele două bile luînd la întîmplare mai întîi una din ele şi apoi pe cea de a doua. Dacă notăm cu A evenimentul care se produce atunci cînd prima bilă extrasă este de culoare albă, iar cu B evenimentul care se produce atunci cînd cea de a doua bilă extrasă este de culoare albă, este clar că producerea evenimentului B este influenţată de producerea evenimentului A, deoarece atunci cînd se produce evenimentul A evenimentul B nu se poate produce.

Deci evenimentele A, B sînt dependente. În cazul cînd prima bilă extrasă se reintroduce în urnă şi apoi se extrage din nou, la întîmplare, o bilă din această urnă, notînd cu A_1 evenimentul care se produce atunci cînd prima bilă extrasă este de culoare albă, iar cu A_2 evenimentul care se produce atunci cînd cea de a doua bilă extrasă este de culoare albă, se poate spune că evenimentele A_1, A_2 sînt independente, producerea oricăruia dintre aceste două evenimente nefiind influenţată în nici un mod de producerea celuilalt.

Această modalitate de a defini independenţa a două evenimente este însă neconvenabilă din mai multe motive, asupra cărora nu vom insista. Remarcăm totuşi faptul că, în problemele concrete, ceea ce interesează în mod deosebit este probabilitatea producerii evenimentelor cu care se lucrează şi atunci ar fi natural să se spună că două evenimente A, B sînt independente atunci cînd probabilitatea producerii oricăruia dintre acestea nu este influenţată de producerea celuilalt eveniment, caz în care evenimentele A, B se mai numesc şi *statistic independente* sau pe scurt *independente*. Sîntem astfel conduşi la următoarea definiţie:

Definiţia 7.4.6.

Fie A, B două evenimente dintr-un cîmp de probabilitate $(\Omega, K, \mathcal{P})$, astfel încît $P(B) \neq 0$. Vom spune că evenimentele A, B sînt independente dacă $P(A|B) = P(A)$, unde prin $P(A|B)$ se înţelege probabilitatea evenimentului A condiţionată de evenimentul B.

Cu alte cuvinte, evenimente A, B sînt statistic independente atunci cînd măsura şanselor de producere ale evenimentului A nu depinde de faptul dacă evenimentul B se produce sau nu. În exemplul ce urmează vom considera două evenimente, corespunzătoare unei experienţe concrete, care în mod intuitiv pot fi considerate independente, (producerea oricăruia dintre ele nu este influenţată de producerea celuilalt), dar care nu sînt statistic independente, prin aceasta justificîndu-se faptul că independenţa statistică a evenimentelor este mai restrictivă decît independenţa intuitivă.

Exemplul 7.4.9.

Să presupunem că dispunem de trei discuri, unul colorat pe ambele feţe în alb, al doilea colorat pe ambele feţe în negru, iar cel de al treilea avînd una

din fețe colorată în alb și cealaltă în negru. Se alege la întîmplare unul dintre aceste discuri, în urma alegerii avînd posibilitatea să observăm culoarea unei singure fețe a discului ales, (de exemplu discul este așezat pe o masă și avem posibilitatea să observăm numai fața superioară a discului ales). Să notăm cu A evenimentul care se produce atunci cînd fața vizibilă a discului este de culoare albă, iar cu B evenimentul care se produce atunci cînd fața invizibilă a discului ales este de culoare neagră. Dacă alegerea se face astfel încît oricare dintre cele șase fețe ale celor trei discuri, din care se alege unul, are aceleași șanse de a apare vizibilă experimentatorului atunci rezultatul experienței poate fi descris prin șase evenimente elementare egal posibile, fiecare dintre acestea corespunzînd uneia din cele șase fețe ale celor trei discuri. Considerînd acum drept măsură a șanselor de realizare ale unui eveniment probabilitatea definită în mod clasic se constată imediat că $P(A)=\frac{1}{2}$, $P(B)=\frac{1}{2}$ iar $P(A\cap B)=\frac{1}{6}$, deoarece în trei dintre cele șase cazuri egal posibile fața vizibilă este albă, tot în trei dintre cele șase cazuri egal posibile fața invizibilă este de culoare neagră și există un singur caz în care fața vizibilă este de culoare albă și cea invizibilă de culoare neagră. Să observăm că nu s-ar putea spune că producerea evenimentului A este influențată în vreun fel de producerea evenimentului B sau invers, culoarea feței vizibile a discului ales fiind determinată numai de modul în care se face alegerea dintre cele trei discuri și în nici un caz de culoarea feței invizibile a acestui disc, (dacă fața invizibilă este de culoare albă atunci fața vizibilă este de culoare albă sau neagră după cum discul ales este cel cu ambele fețe colorate în alb sau cel cu una colorată în alb iar cealaltă colorată în negru). Totuși, ținînd cont de probabilitățile evenimentelor $A, B, A\cap B$, determinate mai sus și de definiția probabilității condiționate, se constată că evenimentele A, B nu sînt statistic independente deoarece $P(A|B)=\frac{1}{3}\neq\frac{1}{2}=P(A)$, situație datorată faptului că șansele de producere ale evenimentului A sînt influențate de producerea evenimentului B.

În cele ce urmează prin evenimente independente vom înțelege numai evenimente statistic independente. Avînd în vedere definiția probabilității condiționate și definiția independenței a două evenimnete putem constata imediat că următoarele două proprietăți sînt adevărate:

1. Dacă evenimentele A, B sînt independente atunci $P(A \cap B) = P(A)\,P(B)$

2. Dacă evenimentele A, B sînt astfel încît $P(A) \neq 0$, $P(B) \neq 0$, iar $P(A \cap B) = P(A)\,P(B)$ atunci $P(A|B) = P(A)$ şi $P(B|A) = P(B)$, adică evenimentele A, B sînt independente.

Demonstraţie: Pentru justificarea primei proprietăţi este suficient să observăm că din faptul că $P(A|B) = P(A)$ urmează:

$$\frac{P(A \cap B)}{P(B)} = P(A)$$

relaţie care este echivalentă cu $P(A \cap B) = P(A)\,P(B)$. Cea de a doua proprietate se justifică în mod cu totul asemănător, condiţiile $P(A) \neq 0$, $P(B) \neq 0$ asigurîndu-se numai faptul că există probabilităţile condiţionate $P(A|B)$ şi $P(B|A)$.

Vom defini acum independenţa unui număr oarecare de evenimente dintr-un cîmp de probabilitate.

Definiţia 7.4.7.

Se spune că evenimentele $A_1, \ldots, A_n$, dintr-un cîmp de probabilitate (Ω, K, P), sînt mutual independente dacă oricare ar fi k dintre aceste evenimente, $A_{i_1}, \ldots, A_{i_k}$, $1 \leq k \leq n-1$, şi $A_j \neq A_{i_1}, \ldots, A_{i_k}$ atunci:

$$P(A_j | A_{i_1} \cap \ldots \cap A_{i_k}) = P(A_j)$$

cu condiţia ca aceste probabilităţi condiţionate să aibă sens. Astfel spus, evenimentele $A_1, \ldots, A_n$ sînt independente atunci cînd măsura şanselor de producere a oricăruia dintre acestea nu este influenţată de producerea simultană a oricîte dintre celelalte evenimente considerate. Folosind metoda inducţiei matematice (vezi exerciţiul 8) se poate arăta cu uşurinţă că, dacă evenimentele $A_1, \ldots, A_n$ sînt mutual independente şi $\{i_1, \ldots, i_k\} \subset \{1, \ldots, n\}$ atunci probabilitatea intersecţiei evenimentelor $A_{i_1}, \ldots, A_{i_k}$ este egală cu produsul probabilităţilor acestor evenimente, adică:

$$P(A_{i_1} \cap \ldots \cap A_{i_k}) = P(A_{i_1}) \ldots P(A_{i_k})$$

Exerciţii:

1. Să se arate că, dacă două evenimente, A, B, dintr-un cîmp de propabilitate (Ω, K, P), sînt astfel încît $P(B) = 0$ atunci $P(A \cap B) = 0$.

2. Fie (Ω, K, P) un cîmp de probabilitate şi $A_1, \ldots, A_n$ n evenimente oarecare din acest cîmp de probabilitate. Să se arată că au loc relaţiile:

a) $P(\bigcup_{i=1}^{n} A_i) \leq \sum_{i=1}^{n} P(A_i)$

b) $P(\bigcap_{i=1}^{n} A_i) \geq \sum_{i=1}^{n} P(A_i) - (n-1)$ (inegalitatea lui Boole).

3. Fie evenimentele A, B_1, B_2 dintr-un cîmp de probabilitate (Ω, K, P), astfel încît $P(B_1) \neq 0$, $P(B_2) \neq 0$, $B_1 \cap B_2 = \Phi$. Să se arate că $P(A|B_1 \cup B_2) \leq P(A|B_1) + P(A|B_2)$.

4. Fie A, B două evenimente independente dintr-un cîmp de probabilitate (Ω, K, P). Să se arate că $P(\bar{A} \cap B) = P(\bar{A})P(B)$, $P(A \cap \bar{B}) = P(A)P(\bar{B})$, $P(\bar{A} \cap \bar{B}) = P(\bar{A})P(\bar{B})$.

5. Fie evenimentele A, B astfel încît $P(A \cap B) = \frac{1}{8}$, $P(A \cap \bar{B}) = \frac{7}{12}$, $P'(\bar{A} \cap B) = \frac{1}{4}$. Să se determine $P(A)$, $P(B)$, $P(A|B)$, $P(A \cup B)$, $P(\bar{A} \cap \bar{B})$.

6. Dacă $P(A) = 0{,}5$, $P(B) = 0{,}4$ iar $P(A \cap B) = 0{,}2$ se poate spune că evenimentele A, B sînt independente?

7. Fie A, B două evenimente dintr-un cîmp de probabilitate (Ω, K, P) astfel încît $P(A) = \frac{1}{2}$, $P(B) = \frac{1}{2}$, $P(A \cup B) = \frac{2}{3}$. Evenimentele A, B sînt independente?

8. Să se arate că atunci cînd evenimentele $A_1, \ldots, A_n$ sînt mutual independente au loc relaţiile:

$$P(A_{i_1} \cap \ldots \cap A_{i_k}) = P(A_{i_1}) \ldots P(A_{i_k}),\ (\forall)\{i_1, \ldots, i_k\} \subset \{1, \ldots, n\}$$

$$P(A_1 \cap \ldots \cap A_n) = P(A_1)\, P(A_2) \ldots P(A_n)$$

Indicaţie: Se ţine cont de relaţia :

$$P(A_{i_1} \cap \ldots \cap A_{i_k}) = P(A_{i_1} \cap \ldots \cap A_{i_{k-1}})\, P(A_{i_k} | A_{i_1} \cap \ldots \cap A_{i_{k-1}})$$

9. Într-o urnă sînt 5 bile albe şi 5 bile negre. Din această urnă se extrag cuccesiv 3 bile, fără ca bila extrasă să fie reintrodusă în urnă. Fie A evenimentul care se produce atunci cînd măcar una dintre primele două bile extrase este de culoare albă, iar B evenimentul care se produce atunci cînd cea de a treia bilă extrasă este de culoare albă.

Să se determine $\mathcal{P}(A)$, $\mathcal{P}(B)$, $\mathcal{P}(B|A)$, $\mathcal{P}(A|B)$, $\mathcal{P}(A \cap B)$, $\mathcal{P}(A \cup B)$, folosind definiţia clasică a probabilităţii.

Formule de calcul cu probabilităţi

Pentru calculul probabilităţilor unor evenimente se utilizează uneori diferite formule care simplifică determinarea acestor probabilităţi. În cele ce urmează vom prezenta unele dintre aceste formule.

Formula de adunare a probabilităţilor

Fie $(\Omega, K, \mathcal{P})$ un cîmp de probabilitate şi $A_1, \ldots, A_n$, n evenimente ale acestui cîmp de probabilitate. Vom dovedi că are loc relaţia:

$$P\left(\bigcup_{i=1}^{n}\right) = \sum_{i=1}^{n} P(A_i) - \sum_{i<j} P(A_i \cap A_j) + \ldots + \qquad 7.4.1$$

$$+ (-1)^{k-1} \sum_{i_1 < i_2 < \ldots < i_k} P(A_{i_1} \cap \ldots \cap A_{i_k} + (-1)^{n-1} P(A_1 \cap \ldots \cap A_n)$$

numită formula de adunare a probabilităţilor. Faptul că această formulă este adevărată se justifică folosind metoda inducţiei matematice. Deoarece

demonstrația în cazul general nu necesită raționamente complicate, vom arăta că formula este adevărată pentru $n = 2,3$. Într-adevăr, dacă $n = 2$ atunci relația (7.4.1) devine:

$$P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 \cap A_2)$$

adevărată conform proprietății 5 dovedite anterior. Dacă $n = 3$ atunci relația (7.3.1) devine:

$$P(A_1 \cup A_2 \cup A_3) = P(A_1) + P(A_2) + P(A_3) - P(A_1 \cap A_2) - P(A_1 \cap A_3) - P(A_2 \cap A_3) + P(A_1 \cap A_2 \cap A_3)$$

și în continuare vom dovedi că aceasta este adevărată.

În baza rezultatelor de pînă acum putem scrie succesiv:

$$P(A_1 \cup A_2 \cup A_3) = P((A_1 \cup A_2) \cup A_3) = P(A_1 \cup A_2) +$$

$$+ P(A_3) - P((A_1 \cup A_2) \cap A\text{-}3) = P(A_1 \cup A_2) + P(A_3) +$$

$$+ P((A_1 \cap A_3) \cup (A_2 \cap A_3)) = P(A_1) + P(A_2) -$$

$$- P(A_1 \cap A_2) + P(A_3) - P(A_1 \cap A_3) - P(A_2 \cap A_3) +$$

$$+ P((A_1 \cap A_3) \cap (A_2 \cap A_3))$$

și deoarece $(A_1 \cap A_3) \cap (A_2 \cap A_3) = A_1 \cap A_2 \cap A_3$ urmează că relația (7.4.1) este dovedită și pentru $n = 3$. În mod cu totul asemănător se arată că formula (7.4.1) este adevărată pentru orice n.

Exemplu 7.4.10.

Să presupunem că se aruncă un zar de trei ori consecutiv și vrem să determinăm probabilitatea ca măcar în una dintre aruncări să se obțină fața marcată cu un punct.

Fie A evenimentul care se produce atunci cînd după măcar una dintre cele trei aruncări se obține fața marcată cu un punct și fie A_1, A_2, A_3 evenimentele care se produc atunci cînd apare fața marcată cu un punct după prima, a doua și respectiv a treia aruncare. Este clar că putem scrie $A = A_1 \cup A_2 \cup A_3$ și deci $P(A) = P(A_1 \cup A_2 \cup A_3)$. Deoarece:

$$P(A_1 \cup A_2 \cup A_3) = P(A_1) + P(A_2) + P(A_3) - P(A_1 \cap A_2) -$$

$$- P(A_1 \cap A_3) - P(A_2 \cap A_3) + P(A_1 \cap A_2 \cap A_3)$$

pentru determinarea probabilităţii evenimentului A este suficient să determinăm probabilităţile din membrul al doilea al ultimei egalităţi. Avînd în vedere că evenimentele A_1, A_2, A_3 sînt independente, iar $P(A_1) = P(A_2) = P(A_3) = 1/6$, urmează că:

$$P(A_1 \cap A_2) = P(A_1 \cap A_3) = P(A_2 \cap A_3) = \frac{1}{36}, \; P(A_1 \cap A_2 \cap A_3) = \frac{1}{216}$$

de unde rezultă că $\mathcal{P}(A) = \frac{91}{216}$.

Formula de înmulţire a probabilităţilor

Dacă (Ω, K, P) este un cîmp de probabilitate, iar $(A_1, \ldots, A_n \in K, \; n \geq 2)$, sînt evenimente astfel încît $P(A_1 \cap \ldots \cap A_{i-1}) \neq 0$, iar $i = 2, \ldots, n$, atunci are loc relaţia (7.4.2):

$$P(A_1 \cap \ldots \cap A_n) = P(A_1)\, P(A_2|A_1)\, P(A_3|A_1 \cap A_2) \ldots \qquad 7.4.2$$

$$\ldots P(A_n|A_1 \cap \ldots \cap A_{n-1})$$

numită formula de înmulţire a probabilităţilor. Justificarea acestei formule se poate face prin inducţie matematică, după numărul n al evenimentelor considerate. Să observăm mai întîi că faptul că $\mathcal{P}(A_1 \cap \ldots \cap A_{i-1}) \neq 0$, $i = 2, \ldots, n$, avem asigurată existenţa probabilităţilor condiţionate din membrul al doilea al relaţiei (7.4.2).

Pentru $n = 2$ relaţia (7.4.2) devine:

$$P(A_1 \cap A_2) = P(A_1)\, P(A_2|A_1)$$

adevărată pe baza definiţiei probabilităţii condiţionate. Să presupunem acum că relaţia (7.4.2) este adevărată pentru intersecţia a cel mult n evenimente,

$n \geq 2$, şi să dovedim că este adevărată şi pentru cazul cînd se consideră intersecţia a $n+1$ evenimente.

Pe baza proprietăţilor probabilităţii, dovedite pînă acum, şi a ipotezei inductive, putem scrie succesiv:

$$P(A_1 \cap \ldots \cap A_{n+1}) = P((A_1 \cap \ldots \cap A_n) \cap A_{n+1}) =$$

$$= P(A_1 \cap \ldots \cap A_n)\, P(A_{n+1} | A_1 \cap \ldots \cap A_n) =$$

$$= P(A_1)\, P(A_2|A_1) \ldots P(A_n|A_1 \cap \ldots \cap A_{n-1})\, P(A_{n+1}|A_1 \cap \ldots \cap A_n)$$

ceea ce arată că formula este adevărată şi pentru intersecţia a $n+1$ evenimente. Prin urmare formula (7.4.2) este adevărată pentru orice $n \in \mathbb{N}$, $n \geq 2$.

Notăm că, atunci cînd evenimentele $A_1, \ldots, A_n$ sînt mutual independente, relaţia (7.4.2) devine $P(A_1 \cap \ldots \cap A_n) = P(A_1) \ldots P(A_n)$, (vezi exerciţiul 8 de mai înainte).

Exemplul 7.4.11.

Într-o urnă se află cinci bile dintre care trei sînt albe şi două sînt negre. Din urnă se extrag succesiv trei bile, fără ca bila extrasă să fie reintrodusă în urnă. Să se determine probabilitatea ca toate cele trei bile să fie de culoare albă.

Fie A evenimentul care se produce atunci cînd toate cele trei extrase sînt albe, iar A_1, A_2, A_3 evenimentele la care se produc atunci cînd prima, a doua şi respectiv a treia bilă extrasă este de culoare albă. Putem scrie atunci $A = A_1 \cap A_2 \cap A_3$ şi deci:

$P(A) = P(A_1 \cap A_2 \cap A_3) = P(A_1)\, P(A_2|A_1)\, P(A_3|A_1 \cap A_2)$. Dar $\mathcal{P}(A_1) = \frac{3}{5}$, $\mathcal{P}(A_2|A_1) = \frac{2}{4}$, $\mathcal{P}(A_3|A_1 \cap A_2) = \frac{1}{3}$ şi atunci $\mathcal{P}(A) = \frac{3}{5} \cdot \frac{2}{4} \cdot \frac{1}{3} = \frac{1}{10}$.

(Pentru a determina $\mathcal{P}(A_2|A_1)$ raţionăm în felul următor: dacă se produce evenimentul A_1, atunci înainte de a efectua cea de a doua extragere în urnă se află 4 bile, dintre care numai două sînt albe, deoarece evenimentul A_1 presupune că prima bilă extrasă este de culoare albă).

Înainte de a trece la prezentarea următoarelor formule de calcul cu propbabilităţi este convenabil să introducem o noţiune nouă şi anume noţiunea de *sistem complet de evenimente*.

Definiţia 7.4.8.

Fie (Ω, K, P) un cîmp de probabilitate şi fie evenimentele $A_1,\ldots,A_n$, din acest cîmp de probabilitate. Spunem că evenimentele $A_1,\ldots,A_n$ formează un *sistem complet de evenimente* dacă sînt îndeplinite condiţiile:

1. $P(A_i) \neq 0$, $(\forall)\, i \in \{1,\ldots,n\}$
2. $A_i \cap A_j = \Phi$, $(\forall)\, i, j \in \{1,\ldots,n\}$, $i \neq j$
3. $A_1 \cup \ldots \cup A_n = \Omega$

Altfel spus, evenimentele $A_1,\ldots,A_n$ formează un sistem complet de evenimente dacă:

1) probabilitatea fiecăruia din ele este nenulă,
2) sînt incompatibile două cîte două şi
3) reuniunea tuturor acestor evenimente este evenimentul sigur.

Formula probabilităţii totale

Fie (Ω, K, P) un cîmp de probabilitate şi $A_1,\ldots,A_n$, n evenimente astfel încît acestea formează un sistem complet de evenimente pentru cîmpul de probabilitate (Ω, K, P).

În acest caz, dacă A este un eveniment oarecare, $A \in K$, atunci are loc relaţia:

$$P(A) = P(A_1)\,P(A|A_1) + P(A_2)\,P(A|A_2) + \ldots + P(A_n)\,P(A|A_n) \qquad 7.4.3$$

numită *formula probabilităţii totale*.

Într-adevăr, deoarece evenimentele A_i, $i = 1,\ldots,n$, îndeplinesc comdiţiile 1) - 3) din definiţia 7.4.8 urmează că putem scrie succesiv:

$$A = A \cap \Omega = A \cap (\bigcup_{i=1}^{n} A_i) = \bigcup_{i=1}^{n} (A \cap A_i)$$

şi deci:

$$P(A) = P(\bigcup_{i=1}^{n} (A \cap A_i))$$

Deoarece $A_i \cap A_j = \Phi$, $(\forall)\, i \neq j$, rezultă că $(A \cap A_i) \cap (A \cap A_j) = \Phi$, deci $(\forall)\, i \neq j$ şi prin urmare:

$$\mathcal{P}(\bigcup_{i=1}^{n} (A \cap A_i)) = \sum_{i=1}^{n} \mathcal{P}(A \cap A_i) = \sum_{i=1}^{n} \mathcal{P}(A_i)\,\mathcal{P}(A|A_i)$$

În concluzie, putem scrie:

$$P(A) = \sum_{i=1}^{n} P(A_i)\, P(A|A_i)$$

ceea ce arată că relaţia (7.4.3) este adevărată.

Formula probabilităţii totale se foloseşte pentru a calcula probabilitatea unui eveniment A, a cărui realizare poate fi considerată ca rezultat al unei experienţe ce ar putea fi descompusă în două etape. În prima etapă rezultatul constă în producerea unui eveniment dintr-un sistem complet de evenimente, iar în a doua etapă rezultatul constă în eventuala realizare a evenimentului A, şansele de producere ale acestuia fiind influenţate de rezultatul obţinut în prima etapă.

Exemplul 7.4.12.

Fie două urne U_1, U_2 fiecare conţinînd cîte două bile albe, urna U_1 conţinînd două bile negre, iar U_2 3 bile negre. Din una din urne, aleasă la întîmplare, se extrage o bilă interesîndu-ne numai culoarea acestei bile. Să determinăm probabilitatea ca bila obţinută să fie de culoare albă.

Să notăm cu A_1 evenimentul care se produce atunci cînd bila obţinută, conform procedeul descris, este extrasă din urna U_1, iar A_2 evenimentul care se produce atunci bila obţinută ar fi extrasă din urna U_2. Deoarece pentru extracţia acestei bile se extrage la întîmplare una dintre cele două urne, urmează că $P(A_1) = \frac{1}{2}$, $P(A_2) = \frac{1}{2}$. Pe de altă parte se constată imediat că: $A_1 \cap A_2 = \Phi$, $A_1 \cup A_2 = \Omega$, unde, ca deobicei, Φ notează evenimentul imposibil, iar Ω evenimentul sigur. Notînd acum prin A evenimentul care se produce atunci cînd bila obţinută este de culoare albă, pentru calculul probabilităţii evenimentului A se poate folosi formula probablităţii totale, pentru $n = 2$, În cazul nostru putem scrie:

$$P(A) = P(A_1)\, P(A|A_1) + P(A_2)\, P(A|A_2)$$

Pentru a determina în acest mod probabilitatea evenimentului A este necesar să mai determinăm doar probabilităţile condiţionate $P(A|A_1)$ şi $P(A|A_2)$. Dar, în ipoteza că se produce evenimentul A_1, urmează că extracţia se face din urna U_1 ce conţine spuă bile albe şi două bile negre, şi atunci probabilitatea ca să se producă şi evenimentul A este 2/4, adică $P(A|A_1) = \frac{1}{2}$.

În mod similar se constată că $P(A|A_2) = \frac{2}{5}$. Deci:

$$P(A) = \frac{1}{2} \cdot \frac{1}{2} + \frac{1}{2} \cdot \frac{2}{5} = 0,45$$

Formula lui Bayes (Formula probabilităţii cauzelor)

Formula pe care o vom prezenta în continuare, este cunoscută sub denumirea de formula, sau teorema lui Bayes. Fără a avea o semnificaţie deosebită în teoria probabilităţilor, este fundamentală pentru un capitol importat al teoriei deciziei, numit teoria Bayesiană a deciziilor.

Fie (Ω, K, P) un cîmp de probabilitate şi evenimentele $A_1, A_2, \ldots, A_n \in K$, astfel încît acestea formează un sistem complet de evenimente pentru (Ω, K, P).

În aceste condiţii, dacă A este un eveniment oarecare al cîmpului de evenimente dat, cu proprietatea că $P(A) \neq 0$, atunci pentru oricare k, $1 \leq k \leq n$, are loc relaţia:

$$P(A_k|A) = \frac{P(A_k)\,P(A|A_k)}{\sum_{i=1}^{n} P(A_i)\,P(A|A_i)} \qquad 7.4.4$$

Această formulă mai poartă denumirea şi de formula probabilităţii, una din explicaţii putînd fi următoarea: Să presupunem că un eveniment, A, se poate produce numai dacă se produce unul dintre evenimentele $A_1, \ldots, A_n$, incompatibile două cîte două evenimentele, $A_1, \ldots, A_n$ pot fi numite „cauze"

posibile ale producerii evenimentului dat A, atunci cînd $\bigcup_{i=1}^{n} A_i = \Omega$, adică măcar una din cauze se produce. Dacă la un moment dat se constată că evenimentul A s-a produs, este interesant de ştiut care este probabilitatea ca producerea evenimentului A să se datorească unei cauze precizate. Deoarece aceste cauze $A_1, \ldots, A_n$ sînt incompatibile două cîte două, este evident că nu se poate să se producă simultan două din acestea, şi atunci probabilitatea căutată poate fi exprimată prin probabilitatea condiţionată $P(A_k|A)$, adică probabilitatea ca să se producă o anumită cauză A_k, în ipoteza că s-a realizat evenimentul A (s-a observat „efectul" A). Pentru determinarea acestei probabilităţi ar putea fi folosită atunci formula (7.4.4). În teoria Bayesiană a deciziilor, probabilitatea $P(A_k)$, $1 \le k \le n$, se mai numeşte *probabilitatea a priori* a cauzei A_k, iar probabilitatea condiţionată $P(A_k/A)$, $1 \le k \le n$, *probabilitatea a-posteriori* a cauzei A_k, aceste denumiri justificîndu-se prin cele spuse mai sus.

Să demonstrăm acum că formula (7.4.4) este adevărată.

Fie A_k, $1 \le k \le n$, unul dintre evenimentele $A_1, \ldots, A_n$. Ţinînd cont de definiţia probabilităţii condiţionate şi de faptul că $P(A) \neq 0$, putem scrie:

$$P(A_k|A) = \frac{P(A_k \cap A)}{P(A)}$$

Dar $P(A_k \cap A) = P(A_k)\,P(A|A_k)$ şi, deoarece evenimentele A_i, $i = 1, \ldots, n$ formează un sistem complet de evenimente, conform formulei probabilităţii totale, are loc şi relaţia:

$$P(A) = \sum_{i=1}^{n} P(A_i)\,P(A|A_i)$$

Deci:

$$P(A_k|A) = \frac{P(A_k)\,P(A|A_k)}{\sum_{i=1}^{n} P(A_i)\,P(A|A_i)}$$

ceea ce demonstrează formula (7.4.4).

Exemplul 7.4.13.

Să considerăm trei urne U_1, U_2, U_3, două dintre acestea conţinînd cîte o bilă albă şi una neagră, iar cea de a treia urnă două bile albe şi una neagră. Din una din urne, luată la întîmplare, se extrage la întîmplare o bilă şi se constată că aceasta este de culoare albă. Ne interesează probabilitatea ca bila extrasă să provină din una din urnele în care se află o bilă albă şi una neagră.

Să notăm că A evenimentul care se produce cînd bila extrasă este de culoare albă, (culoarea bilei extrase este stabilită numai după ce aceasta se extrage din una din cele trei urne), iar cu A_1, A_2 evenimentele care se produc atunci cînd bila extrasă provine din urna în care se află două bile albe şi una neagră, respectiv din una din urnele în care se află o bilă albă şi una neagră. Este clar că $A_1 \cap A_2 = \Phi$ şi $A_1 \cup A_2 = \Omega$. Pe de altă parte $P(A_1) = \frac{1}{3}$, $P(A_2) = \frac{2}{3}$. Probabilitatea care ne interesează este probabilitatea condiţionată $P(A_2|A)$, care poate fi determinată folosind formula lui Bayes, pentru $n = 2$, adică:

$$P(A_2|A) = \frac{P(A_2)\,P(A|A_2)}{P(A_1)\,P(A|A_1) + P(A_2)\,P(A|A_2)}$$

Dar $P(A|A_2) = \frac{1}{2}$, $P(A|A_1) = \frac{2}{3}$ şi atunci, făcînd înlocuirile corespunzătoare în relaţia de mai sus obţinem $P(A_2|A) = 0{,}6$.

Exemplul 7.4.14.

Pentru a se obţine informaţii relativ la nivelul cererii pentru un anumit tip de produs, se decide efectuarea unui sondaj în rîndul consumatorilor. Pentru perioada de timp care interesează, se ajunge la concluzia că ar fi suficient să se ştie dacă cererea pentru acest tip de produs se află între 10 şi 20 unităţi, sau între 21 şi 40 unităţi, sau mai mare decît 41 unităţi. În mod convenţional vom spune că cererea este *mare* dacă depăşeşte 40 unităţi de produs, *medie* dacă este între 21 şi 40 unităţi de produs şi mică dacă este între 10 şi 20 unităţi de produs. Din modul în care este organizat sondajul, cît şi din experienţa enterioară, se ştie că în cazul în care cererea este în realitate mică, rezultatul sondajului va indica despre aceasta că este mică cu probabilitatea 0,8 şi că este medie cu probabilita-

tea 0,2. De asemeni, dacă în realitate cererea este medie, sondajul va indica aceasta cu probabilitatea 0,7, cu probabilitatea 0,2 indicînd că cererea este ridicată, iar cu probabilitatea 0,1 indicînd că cererea este mică. În sfîrşit atunci cînd cererea este în realitate mare, sondajul va indica aceasta cu probabilitatea 0,8, iar cu probabilitatea 0,2 va indica un nivel mediu al cererii. Să mai presupunem că, tot din experienţa anterioară, se ştie că pentru intervale de timp de tipul celui considerat, cererea este mică cu probabilitatea 0,3, este medie cu probabilitatea 0,5 şi mare cu probabilitatea 0,2.

După efectuarea sondajului se constată că acesta indică faptul că, pentru intervalul de timp respectiv, cererea este medie şi vrem să ştim care este probabilitatea ca în realitate cererea să fie de nivel mediu.

În acest scop să notăm prin A_1, A_2, A_3 evenimentele care se produc atunci cînd cererea din produsul considerat este mică, medie şi mare, respectiv, iar cu A evenimentul care se produce atunci cînd rezultatul sondajului ce urmează a fi efectuat indică faptul că cererea este de nivel mediu. Conform celor spuse mai înainte, din datele disponibile pînă la efectuarea sondajului se acceptă că $P(A_1) = 0,3$, $P(A_2) = 0,5$, $P(A_3) = 0,2$, acestea fiind probabilităţile evenimentelor A_1, A_2, A_3 înainte de a obţine informaţiile furnizate de sondaj, numite şi „probabilităţi a-priori" ale evenimentelor A_1, A_2, A_3. De asemeni, se mai cunosc probabilităţile $P(A|A_1) = 0,2$, $P(A|A_2) = 0,7$, $P(A|A_3) = 0,2$, deoarece, în ipoteza că cererea este mică sau mare, sondajul va indica despre aceasta că este medie cu probabilitatea 0,2, iar în cazul cînd cererea este medie, sondajul va indica aceasta cu probabilitatea 0,7. Deoarece evenimentele A_1, A_2, A_3 se exclud reciproc şi $A_1 \cup A_2 \cup A_3 = \Omega$, rezultă că probabilitatea condiţionată $P(A_2|A)$, care ne interesează, se poate determina folosind formula (7.4.4) pentru $n = 3$, adică:

$$P(A_2|A) = \frac{P(A_2)\,P(A|A_2)}{P(A_1)\,P(A|A_1) + P(A_2)\,P(A|A_2) + P(A_3)\,P(A|A_3)}$$

După înlocuiri obţinem $P(A_2|A) = \frac{7}{9} \approx 0,78$. Deci probabilitatea ca rezultatul sondajului efectuat să indice situaţia reală a cererii este $\frac{7}{9}$, aceasta

purtînd denumirea de „probabilitate à-posteriori" a unei cereri de nivel mediu, deoarece a fost determinată după ce s-au obtinut informaţii suplimentare privind cererea, informaţii furnizate de sondajul respectiv.

În mod similar se pot determina probabilitatea à-posteriori ale unei cereri de nivel scăzut sau ridicat obţinîndu-se $P(A_1|A) = 0,13$, $P(A_3|A) = 0,09$.

Scheme probabilistice clasice

În practică se întîlnesc situaţii în care pentru a se studia caracteristicile unei colectivităţi se fac experienţe, sau sondaje, conform unor scheme teoretice bine precizate. În cele ce urmează vom prezenta trei astfel de scheme şi anume: schema bile nerevenite, schema bilei revenite şi schema urnelor Poisson.

Schema bilei nerevenite

Să considerăm că într-o urnă, U, se află a bile albe şi b bile negre. Din această urnă se extrag la întîmplare n bile, una cîte una, $(n \leq a + b)$, fără ca bila extrasă să fie reintrodusă în urnă. Între bilele extrase se vor afla, în general, atît bile albe cît şi bile negre, numărul bilelor din fiecare culoare nefiind cunoscut decît după ce au fost făcute toate cele n extragerii. Să presupunem că ne-ar interesa ca, între cele n bile ce urmează a fi extrase, un număr dat de bile α, $\alpha \leq a$, să fie de culoare albă. Desigur că, după efectuarea celor n extrageri vom putea constata dacă între cele n bile se află α bile albe şi $n - \alpha$ bile negre, sau nu. Să notăm prin $A(a, b; n, \alpha)$ evenimentul care se produce atunci cînd în urna celor n extrageri vom obţine α bile albe şi $(n - \alpha)$ bile negre şi, folosind definiţia clasică a probabilităţii, ne propunem să determinăm probabilitatea acestui eveniment.

Deoarece după fiecare extragere bila extrasă nu este reintrodusă în urnă, numărul cazurilor egal posibile este acelaşi cu numărul modurilor în care se pot alege n bile din cele $(a + b)$ bile existente în urnă, adică C_{a+b}^{n}. Pe de altă

parte, producerea evenimentului $A(a,b;n,\alpha)$ este echivalentă cu producerea măcar a unuia dintre cazurile egal posibile în care numărul bilelor albe este α şi, prin urmare, numărul cazurilor favorabile producerii evenimentului $A(a,b;n,\alpha)$ este egal cu numărul grupurilor distincte de cîte n bile, ce se pot forma din cele $(a+b)$ bile existente în urnă, care conţin exact α bile albe şi $(n-\alpha)$ bile negre. Datorită faptului că o bilă extrasă nu se reintroduce în urnă, pentru a obţine α bile albe este necesar ca între cele n bile extrase să se afle grupă oarecare de α bile dintre cele a bile albe existente în urnă, numărul acestor grupe fiind C_a^{α}, iar restul de $(n-\alpha)$ bile putînd fi oricare grupă de $(n-\alpha)$ bile negre dintre cele b bile negre existente în urnă, numărul acestor grupe fiind $C_b^{n-\alpha}$. Prin urmare, dintre cele C_{a+b}^{n} grupe de cîte n bile exact $C_a^{\alpha}C_b^{n-\alpha}$ grupe conţin α bile albe şi $(n-\alpha)$ bile negre. Deci probabilitatea evenimentului $A(a,n;n,\alpha)$ este:

$$P(A(a,b;n,\alpha)) = \frac{C_a^{\alpha}C_b^{n-\alpha}}{C_{a+b}^{n}} \qquad 7.4.5$$

Notăm faptul că, dacă $\alpha < a$ sau $(n-\alpha) > b$, atunci evenimentul $A(a,b;n,\alpha)$ este evenimentul imposibil şi deci are probabilitatea egală cu zero.

Observaţia 7.5. Dacă se consideră o urnă în care se află bile de s culori diferite, $(s \geq 2)$, şi notăm cu A evenimentul care se produce atunci cînd extrăgînd la întîmplare n bile, una cîte una fără ca bila extrasă să fie reintrodusă în urnă, între acestea se află α_i bile, de culoarea i, $i = 1,\ldots,s$ se constată că probabilitatea evenimentului A este:

$$P(A) = \frac{C_{a_1}^{\alpha_1}C_{a_2}^{\alpha_2}\ldots C_{a_s}^{\alpha_s}}{C_N^{n}} \qquad 7.4.6$$

unde $a_i, i = 1,\ldots,s$, este numărul de bile de culoarea i existente iniţial în urnă, iar $N = a_1 + a + \ldots + a_s$.

Exemplul 7.4.15.

Să presupunem ca într-o urnă U se află 8 bile de culoare albă şi 10 bile de culoare neagră. Din această urnă urmează a se extrage succesiv 6 bile, fără ca bila extrasă să fie reintrodusă în urnă (fără revenire). Să determinăm probabilitatea ca între cele 6 bile extrase să fie afle exact 4 de culoare albă.

Notînd cu A evenimentul care se produce atunci cînd după extragerea celor 6 bile se constată că 4 dintre ele sînt albe şi 2 sînt negre, deoarece extragerile se fac conform schemei bilei nerevenite, urmează că:

$$P(A) = \frac{C_8^4 C_{10}^2}{C_{18}^6}$$

de unde, după efectuarea calculelor, se obţine $P(A) = \frac{75}{442} \approx 0,17$.

Exemplul 7.4.16.

Dintr-un lot de 10.000 becuri se iau la întîmplare 100 becuri, pentru a se cerceta dacă acestea îndeplinesc condiţiile de calitate necesare. Dacă se ştie că din cele 10.000 de becuri 200 nu îndeplinesc condiţiile de calitate, care este probabilitatea ca între cele 100 de becuri alese în vederea testării să nu se afle becuri care să nu îndeplinească condiţiile de calitate necesare.

Să notăm cu A evenimentul care se produce atunci cînd toate cele 100 becuri alese în vederea testării îndeplinesc condiţiile necesare. Deoarece ceea ce interesează la aceste becuri este îndeplinirea sau nu a condiţiilor de calitate observăm că alegerea şi verificarea celor 100 becuri se face conform schemei bilei nerevenite, dacă se presupune că becurile care îndeplinesc condiţiile necesare se identifică cu bile albe, iar cele care nu le îndeplinesc se identifică cu bile negre. Deci:

$$P(A) = \frac{C_{9800}^{100} C_{200}^0}{C_{10.000}^{100}}$$

Pentru a se obţine o formă explicită a probabilităţii evenimentului A se pot folosi tabele de logaritmi şi se obţine $P(A) \approx 0,15$.

Schema bilei revenite

Fie o urnă U în care se află bile de două culori, de exemplu albe şi negre şi, să presupunem că se cunoaşte probabilitatea evenimentului A, eveniment care se produce atunci cînd luînd la întîmplare o bilă din urnă aceasta este albă. Să notăm cu p probabilitatea evenimentului A, adică $P(A) = p$.

Contrarul evenimentului A se va produce atunci cînd luînd la întîmplare o bilă din urna dată aceasta este de culoare neagră, deci $P(\bar{A}) = 1 - p$.

Dacă din urna U se extrag succesiv n bile, după fiecare extragere reintroducînd bila în urnă, să notăm prin $A(p; n, \alpha)$ evenimentul care se produce atunci cînd în urma celor n extrageri de α ori bila extrasă este albă $0 \le \alpha \le n$, iar de $(n - \alpha)$ ori bila extrasă este neagră. Ne propunem să determinăm probabilitatea evenimentului $A(p; n, \alpha)$.

În acest scop, să notăm cu A_i, $1 \le i \le n$, evenimentul care se produce atunci cînd în extragerea de ordin i se obţine o bilă albă, iar prin $E(i_1, \ldots, i_\alpha)$, $1 \le i_1 < \ldots < i_\alpha \le n$, evenimentul care se produce atunci cînd în extragerile de ordin $i_1, i_2, \ldots, i_\alpha$ se obţin bile albe şi în celelalte $(n - \alpha)$ extrageri se obţin bile negre. Să observăm că se pot construi în total C_n^α evenimente de forma $E(i_1, \ldots, i_\alpha)$, iar pe de altă parte putem scrie:

$$A(p; n, \alpha) = \bigcup E(i_1, \ldots, i_\alpha)$$

reuniunea făcîndu-se după toate modurile distincte de alegere a numerelor $i_1, \ldots, i_\alpha$ astfel încît $1 \le i_1 < \ldots < i_\alpha \le n$.

Totodată din modul în care este definit evenimentul $E(i_1, \ldots, i_\alpha)$ este clar că are loc relaţia:

$$E(i_1, \ldots, i_\alpha) = A_{i_1} \bigcap \ldots \bigcap A_{i_\alpha} \bigcap \left(\bigcap_{i \notin \{i_1, \ldots, i_\alpha\}} \bar{A}_i \right)$$

Deoarece probabilitatea de a obţine o bilă albă în oricare dintre cele n extrageri nu este influenţată în nici-un mod de rezultatul celorlalte extrageri, putem spune că evenimentele A_i, $i = 1, \ldots, n$, sînt mutual independente şi deci

evenimentele $A_{i_1},\ldots,A_{i_\alpha}$, $\bar{A}_i$, $i = \neq i_1,\ldots,i_\alpha$, sînt independente, ceea ce ne permite să scriem relaţia:

$$P(E(i_1,\ldots,i_\alpha)) = P(A_{i_\alpha}) \prod_{1 \neq i_1,\ldots,i_\alpha} P(\bar{A}_i)$$

Observînd acum că probabilitatea obţinerii unei bile albe, în oricare dintre extrageri, este aceeaşi cu probabilitatea evenimentului A, adică $P(A_i) = p$, $P(\bar{A}_i) = 1 - p$, $i = 1,\ldots,n$, rezultă că are loc:

$$P(E(i_1,\ldots,i_\alpha)) = p^\alpha (1-p)^{n-\alpha} \qquad 7.4.7$$

oricare ar fi $i_1,\ldots,i_\alpha$, astfel că $1 \le i_1 < \ldots < i_\alpha \le n$.

Pentru a determina acum probabilitatea evenimentului $A(p;n,\alpha)$ este suficient să observăm că evenimentele $E(i_1,\ldots,i_\alpha)$, a căror reuniune este $A(p;n,\alpha)$, sînt incompatibile două cîte două, ceea ce ne permite să scriem relaţia:

$$P(A(p;n,\alpha)) = \sum_{i \le i_1 < \ldots < i_\alpha \le n} P(E(i_1,\ldots,i_\alpha)) \qquad 7.4.8$$

După cum am mai remarcat, numărul evenimentelor $E(i_1,\ldots,i_\alpha)$, $1 \le i_1 < \ldots < i_\alpha \le n$ este C_n^α şi atunci, pe baza relaţiilor (7.4.7), (7.4.8), urmează:

$$P(A(p;n,\alpha)) = C_n^\alpha p^\alpha (1-p)^{n-\alpha} \qquad 7.4.9$$

Dacă se notează $1 - p = q$, observăm că $P(A(p;n,\alpha))$ este coeficientul lui t^α din dezvoltarea binomului $(pt+q)^n$, fapt pentru care această schemă se mai numeşte şi *schema binomială*.

Observaţia 7.6. Probabilitatea evenimentului $A(p;n,\alpha)$ a fost determinată în cazul în care acest eveniment corespunde unei experienţe care constă în extragerea succesivă a n bile dintr-o urnă în care se află bile albe şi negre. Remarcăm faptul că reintroducerea bilei în urnă după fiecare extracţie nu este semnificativă pentru modul în care am determinat probabilitatea evenimentului $A(p;,n,\alpha)$, semnificativ fiind faptul că probabilitatea de a obţine o bilă albă, în fiecare extragere, rămîne neschimbată (aceasta putîndu-se realiza şi în alt mod, de exemplu

nu prin reintroducerea în urnă a bilei extrase ci a unei alte bile de aceeaşi culoare cu bila extrasă).

Pe de altă parte, în această schemă putem încadra o clasă mai largă de experienţe decît experienţele care se pot reduce la extragerea succesivă a unor bile dintr-o urnă. Să presupunem că o experienţă constă în efectuarea a n probe, consecutiv, în fiecare dintre aceste probe un eveniment A producîndu-se cu probabilitatea p şi contrarul său, $\bar{A}$, cu probabilitatea $q = 1 - p$. Notîndu-se prin $A(p; n, \alpha)$ evenimentul care se produce atunci cînd în urma efectuării celor n probe, de exact α ori, (în α dintre probe), se produce evenimentul A, iar în celelalte $(n - \alpha)$ probe se produce evenimentul $\bar{A}$, în cazul cînd cele n probe sînt *independente*, urmînd aceeaşi cale de mai înainte se constată că probabilitatea evenimentului $A(p; n, \alpha)$ este cea dată de relaţia (7.4.9). Rezultatele anterioare pot fi extinse fără dificultate la cazul unei urne în care se află bile de s culori, $s \geq 2$.

Presupunînd că dintr-o astfel de urnă se extrag succesiv n bile, după fiecare extragere bila extrasă fiind reintrodusă în urnă (extrageri cu revenire), dacă se notează cu A evenimentul care se produce atunci cînd între cele n bile extrase exact n_i bile, $1 \leq i \leq s$, sînt de culoare i, atunci probabilitatea evenimentului A este:

$$P(A) = \frac{n!}{(n_1)!\,(n_2)!\ldots(n_s)!}\, p_1^{n_1} p_2^{n_1} \cdots p_s^{n_s} \qquad 7.4.10$$

unde p_i, $1 \leq i \leq s$, este probabilitatea de a extrage o bilă de culoare i din urna considerată. În mod similar, dacă într-o experienţă care constă din n probe independente, în fiecare dintre aceste probe fiind posibilă producerea unuia dintre s evenimente incompatibile două cîte două, $A_1, A_2, \ldots, A_s$, atunci probabilitatea ca în cele n probe evenimentul A_i, $1 \leq i \leq s$, să se producă de exact n_i ori, $i = 1, 2, \ldots, s$, este dată de relaţia (7.4.10), unde p_i, $1 \leq i \leq s$, este probabilitatea producerii evenimentului A_i în oricare dintre cele n probe $(n_1 + \ldots + n_s = n)$.

Exemplul 7.4.17.

Să presupunem că se cunoaştem că probabilitatea ca un bec de un anumit tip să funcţioneze cel puţin 1000 ore, este 0,8. Dacă se achiziţionează 5 astfel de

becuri, care este probabilitatea ca măcar patru dintre ele să funcţioneze cel puţin 1000 ore, fiind acceptată iptoteza că timpul de funcţionare al oricărui dintre becuri este independent de timpul de funcţionare al celorlalte?

Pentru a răspunde la întrebare, să observăm că dacă notăm cu A evenimentul care se produce atunci cînd un bec luat la întîmplare funcţionează cel puţin 1000 ore, rezultă că $P(A) = 0,8$ şi ne interesează probabilitatea ca din 5 astfel de becuri cel puţin patru să funcţioneze cel puţin 1000 ore. Să notăm cu B evenimentul care se produce, atunci cînd exact 4 dintre cele 5 becuri achiziţionate funcţionează cel puţin 1000 ore, iar cu C evenimente care se produc atunci cînd toate cele 5 becuri funcţionează cel puţin 1000 de ore.

Pentru ca cel puţin 4 dintre cele 5 becuri să funcţioneze măcar 1000 ore trebuie ca măcar unul dintre evenimentele B, C să se producă, adică probabilitatea ca cel puţin 4 dintre cele 5 becuri să funcţioneze măcar 1000 ore este $P(B \cup C)$. Deoarece $B \cap C = \Phi$ urmează că $P(B \cup C) = P(B) + P(C)$ şi, deoarece evenimentul B se produce atunci cînd, luînd 5 becuri, de exact 4 ori se produce evenimentul A, iar evenimentul C se produce atunci cînd în toate cele 5 cazuri se produce evenimentul A, rezultă că $P(B), P(C)$ se pot determina folosind relaţia (7.4.9), adică:

$$P(B) = C_5^4 (0,8)^4 \cdot (0,2)^1,$$
$$P(C) = C_5^5 (0,8)^5 \cdot (0,2)^0,$$

de unde se obţine $P(B) + P(C) = 0,7373$. Deci probabilitatea ca cel puţin 4 dintre cele 5 becuri să funcţioneze măcar 1000 ore este aproximativ egală cu 0,74.

Schema urnelor Poisson

În schema bilei revenite este caracteristic faptul că, la fiecare extragere, probabilitatea obţinerii unei bile de o anumită culoare este aceeaşi. Dacă presupunem că aceste probabilităţi sînt cunoscute, dar diferite de la o extracţie la alta, sîntem conduşi la o nouă schemă, numită schema urnelor Poisson (sau schema Poisson). Fie n urne $U_1, \ldots, U_n$ în care se află bile albe şi negre, p_i, $1 \le i \le n$, fiind probabilitatea de a extrage o bilă albă din urna U_i. Atunci $q_i = 1 - p_i$, $1 \le i \le n$, este evident probabilitatea de a extrage o bilă neagră din urna U_i. Să

considerăm că din fiecare din cele n urne se extrage cîte o bilă şi să notăm cu A evenimentul care se produce atunci cînd între cele n bile extrase se află exact α bile albe. Dacă se notează cu $A_i, 1 \le i \le n$, evenimentul care se produce atunci cînd bila extrasă din urna U_1 este albă, evenimentul A_i pot fi considerate independente şi, descompunînd evenimentul A în mod asemănător descompunerii evenimentului $A(p; n, \alpha)$, definit în cadrul schemei bilei revenite, se poate arăta că probabilitatea evenimentului A, $P(A)$ este egală cu coeficientul lui t^α din dezvoltarea produsului:

$$(p_1 t + q_1) \cdot (p_2 t + q_2) \ldots (p_n t + q_n) \qquad 7.4.11$$

Cu alte cuvinte, dacă:

$$(p_1 t + q_1)(p_2 t + q_2) \ldots (p_n t + q_n) = \sum_{k=o}^{n} a_k t^k$$

atunci $P(A) = a_\alpha$.

Observaţia 7.7. Dacă o experienţă constă în n probe independente, în fiecare dintre aceste probe fiind posibilă producerea unui eveniment A şi, dacă în proba de ordin i evenimentul A se produce cu probabilitatea $p_i, 1 \le i \le n$, atunci probabilitatea ca în cele n probe evenimentul A să se producă de exact α ori, (în α dintre probe să se producă evenimentul A, iar în celelalte $(n - \alpha)$ probe evenimentul $\bar{A}$), este dată deasemeni de coeficientul lui t^α din polinomul (7.4.11).

Exemplul 7.4.18.

Să presupunem că în trei grupe de studenţi, G_1, G_2, G_3, sînt studenţi cu note de cel puţin 7 la o anumită disciplină, în felul următor: în grupa G_1, din 28 studenţi, 21 au note mai mari sau egale cu 7, în grupa G_2, din 30 studenţi, 24 au note de cel puţin 7, iar în grupa G_3, din 27 studenţi, 18 au note de cel puţin 7. Din fiecare dintre aceste grupe se alege la întîmplare cîte un student. Care este probabilitatea ca numai doi dintre cei trei studenţi aleşi să aibă note de cel puţin 7 la disciplina respectivă?

Notînd cu $A_i, i = 1, 2, 3$, evenimentele corespunzătoare alegerii unui

student cu note peste 7 din grupele G_i, $i = 1, 2, 3$, respectiv, iar cu A evenimentul care se produce atunci cînd exact 2 dintre studenţi aleşi au note peste 7 inclusiv rezultă că $P(A_1) = \frac{3}{4}$, $P(A_2) = \frac{3}{5}$, $P(A_3) = \frac{2}{3}$.

Deci probabilitatea evenimentului A, $P(A)$, este coeficientul lui t^2 din polinomul:

$$\left(\frac{3}{4}t + \frac{1}{4}\right)\left(\frac{4}{5}t + \frac{1}{5}\right)\left(\frac{2}{3}t + \frac{1}{3}\right)$$

adică $P(A) = \frac{3}{4}\frac{4}{5}\frac{1}{3} + \frac{3}{4}\frac{2}{3}\frac{1}{5} + \frac{4}{5}\frac{2}{3}\frac{1}{5}$. Efectuînd calculele obţinem: $P(A) \approx 0{,}4$.

Exerciţii:

1. Se aruncă un zar de trei ori consecutiv, în fiecare dintre aruncări fiecare dintre feţe avînd aceleaşi şanse de apariţie. Care este probabilitatea ca măcar în una dintre aruncări să apară faţa marcată cu 6 puncte?

2. Un test constă în 10 întrebări la care trebuie răspuns prin DA sau NU. În cazul cînd răspunsul corect este DA pentru 5 dintre întrebări şi NU pentru celelalte 5 întrebări, care este probabilitatea ca punînd la întîmplare răspunsurile DA, NU, măcar la una dintre întrebări răspunsul să fie corect.

3. Într-o urnă sînt 3 bile albe şi 5 bile negre. Dacă din această se extrag succesiv 4 bile, fără ca bila extrasă să fie reintrodusă în urnă, care este probabilitatea ca prima şi ultima bilă extrasă să fie albe, iar celelalte două negre.

4. Pentru admiterea la cursurile de calificare într-o anumită meserie, candidaţii sînt supuşi unui test de aptitudini. Din experienţa trecută se ştie că 70% dintre persoanele care au fost testate au obţinut calificativul „satisfăcător" în urma testării, iar restul calificativului „nesatisfăcător". Din modul în care este organizat testul se mai ştie că o persoană care are aptitudini pentru meseria respectivă poate trece testul cu propabilitatea 0,8, iar o persoană care în realitate nu are aptitudini pentru această meserie poate trece testul cu probabilitatea 0,3. Avînd la dispoziţie aceste informaţii, să se determine probabilitatea ca un candidat care a obţinut calificativul „satisfăcător" la testarea aptitudinilor să aibă *în realitate* aptitudinile necesare practicării meseriei considerate.

5. Un număr de 100 unităţi dintr-un anumit tip de produs este împărţit în două loturi de cîte 50 unităţi de produs. Dacă între cele 100 unităţi considerate se află 10 unităţi care prezintă defecţiuni, care este probabilitatea ca fiecare din-

tre cele două loturi de cîte 50 unităţi de produs să conţină acelaşi număr de unităţi cu defecţiuni? Dar dacă între cele 100 unităţi de produs se află 9 unităţi cu defecţiuni, care este probabilitatea ca fiecare dintre cele două loturi să conţină cel mult 4 unităţi cu defecţiuni?

6. Într-o secţie a unei întreprinderi se folosesc semifabricate prelucrate în celelalte secţii ale întreprinderii. Dintr-o analiză a modului în care această secţie este aprovizionată cu semifabricate, s-a constatat că 25% din cererile de semifabricate ale acestei secţii nu pot fi satisfăcute la timp. Dacă faptul că o anumită cerere este satisfăcută la timp nu este influenţată de modul în care sînt satisfăcute celelalte cereri, să se determine probabilitatea ca din zece cereri consecutive

a) cel mult trei să nu fie satisfăcute la timp,

b) să nu fie satisfăcute la timp între trei şi cinci cereri, inclusiv,

c) să fie satisfăcute la timp cel puţin două dintre cele 10 cereri.

7. Probabilitatea ca o cantitate dată de N unităţi dintr-un anumit tip de produs să fie vîndută în întregime în decurs de o zi este egală cu 0,5, 0,6, 0,8 în funcţie de care dintre unităţile de desfacere U_1, U_2, U_3, respectiv, asigură distribuţia acestui produs. Dacă într-o anumită zi fiecare dintre cele trei unităţi U_1, U_2, U_3 primesc o cantitate de N unităţi de produs, care este probabilitatea ca la cel mult două dintre acestea cantitatea respectivă să fie epuizată în ziua în care este primită? Pentru calculul acestei probabilităţi, folosind numai datele de mai sus, sînt necesare şi alte ipoteze sau nu?

§ 5. VARIABILE ALEATOARE

Rezultatul unei experienţe este de multe ori convenabil să fie exprimat printr-un număr, sau un ansamblu ordonat de numere. Dacă am lua în consideraţie numai schemele probabilistice prezentate în capitolul precedent, atunci cînd urnele ar conţine numai bile albe şi negre, am putea asocia numărul α evenimentului care se produce atunci cînd din n extrageri, de exact α ori se obţin bile albe, deoarece nu ne interesează care dintre bilelele albe vor fi extrase şi care dintre extrageri vom obţine bile albe. Să observăm că în cazul schemei bilei revenite, de exemplu, dacă urna considerată conţine bile de s culori, $s \geq 2$, evenimentul care se produce atunci cînd între cele n bile extrase se află n_i bile de culoarea i, $1 \leq i \leq s$, putem să-i asociem ansamblul de numere $(n_1, \ldots, n_s)$, dar

pentru identificarea evenimentului este suficient să-i asociem ansamblul de numere $(n_1,\dots,n_{s-1})$ deoarece $n_s = n - (n_1 + \dots + n_{s-1})$. În primul caz numărul α, iar în cel de al doilea caz ansamblul ordonat de numere $(n_1,\dots,n_{s-1})$, nu se cunoaşte decît după ce au efectuat toate cele n extrageri din cadrul experienţei respective, înainte de efectuarea experienţei avînd posibilitatea să precizăm doar faptul că numărul bilelor albe, α, poate fi un număr din mulţimea $\{0,1,\dots,n\}$, pentru primul caz, iar pentru al doilea caz ca numărul bilelor de culoarea i poate fi un număr din mulţimea $\{0,1,\dots,n\}$, ansamblul ordonat $(n_1, n_2, \dots, n_{s-1})$ aflîndu-se în mulţimea $\{(x_1,\dots,x_{s-1}); 0 \le x_i \le n, \sum x_i \le n\}$. De obicei, atunci cînd o anumită mărime poate avea orice valoare dintr-o mulţime cu cel puţin două elemente, această mărime se numeşte *variabilă*. Dacă o variabilă poate lua valori într-o mulţime de numere reale atunci se spune că această variabilă este o mulţime unidimensională, iar în cazul cînd prin valori ale variabilei se înţeleg succesiuni ordonate de numere reale, variabila respectivă se numeşte variabilă multidimensională. Cu alte cuvinte, dacă o variabilă poate lua valori dintr-o mulţime, $D, D \subset R$, această variabilă se spune că este o variabilă unidimensională, iar dacă o variabilă poate lua valori dintr-o mulţime D, $D \subset R^n$, $n \ge 2$ variabila respectivă se spune că este o variabilă multidimensională, mai precis n-dimensională.

Atunci cînd dintr-o urnă cu bile albe şi negre se extrag succesiv n bile şi ne interesează numărul de bile albe obţinute în urma celor n extrageri, atunci, în mod natural, vom asocia acestei experienţe o variabilă care poate lua valori din mulţimea $\{0,1,\dots,n\}$, deci o variabilă unidimensională, valorile acestei variabile reprezintă numărul de bile albe care ar putea fi obţinute după ce se fac cele n extrageri. Dacă în urnă sînt bile de s culori şi extrag succesiv n bile, interesîndu-ne cîte bile din fiecare culoare se vor afla între cele n bile extrase, atunci unei astfel de experienţe îi putem asocia o variabilă care poate lua valori din mulţimea $\{(x_1,\dots,x_{s-1}),\ 0 \le x_i \le n, \sum_{i=1}^{s-1} x_i \le n\}$, deci o variabilă $(s-1)$-dimensională, x_i, $1 \le i \le s-1$, reprezintă numărul bilelor de culoare i obţinute în urma celor n extrageri, (numărul bilelor de culoarea s va fi $n - \sum_{i=1}^{s-1} x_i$). În cazul

cînd o variabilă ia valori dintr-o mulţime dată în funcţie de producerea sau nu a unor evenimente, în mod obişnuit variabila respectivă este numită *variabilă aleatoare*, adică variabilă ale cărei valori sînt influenţate de „întîmplare", sau de fenomene a căror comportare este imprevizibilă. În viaţa de toate zilele astfel de variabile se întîlnesc în mod curent, de exemplu: necesarul de piese de schimb pentru un anumit utilaj într-o perioadă determinată, volumul vînzărilor unui magazin în decurs de o lună, volumul cererii dintr-un anumit tip de produs în decurs de un trimestru, numărul persoanelor care se află într-o staţie de autobuz la un moment dat, numărul copiilor care se nasc în decurs de o săptămînă într-o anumită ţară sau pe întreg globul, momentul apariţiei unei avarii la un anumit utilaj, după ultima revizie, etc.

Variabile aleatoare unidimensionale

Cele spuse mai sus relativ la noţiunea de variabilă aleatoare nu constituie desigur o definiţie în sensul adevărat al cuvîntului. Vom da acum o definiţie exactă a noţiunii de variabilă aleatoare.

Definiţia 7.5.1.

Fie (Ω, K, P) un cîmp de probabilitate şi fie $X: \Omega \to \mathbb{R}$ o funcţie definită pe Ω şi cu valori reale. Vom spune că X este o variabilă aleatoare unidimensională definită peste cîmpul de probabilitate (Ω, K, P) dacă pentru oricare a, $a \in \mathbb{R}$, are loc relaţia:

$$\{\omega / \omega \in \Omega, X(\omega) < a\} \in K \qquad 7.5.1$$

Cu alte cuvinte, o funcţie $X: \Omega \to \mathbb{R}$ este o varibilă aleatoare unidimensională peste cîmpul de probabilitate (Ω, K, P) dacă pentru orice $a, a \in \mathbb{R}$, submulţimea $\{\omega / \omega \in \Omega, X(\omega) < a\}$ a mulţimii Ω este un eveniment al cîmpului de evenimente (Ω, K).

Exemplul 7.5.1.

Fie cîmpul de probabilitate (Ω, K, P) definit $\Omega = \{\omega_1, \omega_2, \omega_3\}$, $K = \{\Phi, \Omega\{\omega_1, \omega_2\}, \{\omega_3\}\}$, $P(\Omega) = 1$, $P(\Phi) = P(\{\omega_1, \omega_2\}) = P(\{\omega_3\}) = 0$. În

acestă situaţie funcţia $X:\Omega\to\mathbb{R}$, definită prin $X(\omega_1)=X(\omega_2)=0, X(\omega_3)=1$ este o variabilă aleatoare peste cîmpul de probabilitate (Ω, K, P) deoarece:

$$\{\omega/\omega\in\Omega, X(\omega)<a\}=\begin{cases}\Phi, a\le 0\\ \{\omega_1,\omega_2\}, 0<a\le 1\\ \{\omega_1,\omega_2,\omega_3\}=\Omega, a>1\end{cases}.$$

ceea ce arată că pentru orice valoare a lui $a\in\mathbb{R}$ are loc relaţia

$$\{\omega/\omega\in\Omega, X(\omega)<a\}\in K$$

Considerînd acelaşi cîmp de probabilitate (Ω, K, P) şi funcţia $Y:\Omega\to\mathbb{R}$, definită prin $Y(\omega_1)=0$, $Y(\omega_2)=1$, $Y(\omega_3)=2$ constată că funcţia Y nu este o variabilă aleatoare peste acest cîmp de probabilitate, deoarece, de exemplu, mulţimea $\{\omega/\omega\in\Omega, Y(\omega)<1\}=\{\omega_1\}\notin K$ şi deci relaţia (7.5.1) nu este verificată de funcţia Y cel puţin pentru cazul cînd $a=1$.

Exemplul 7.5.2.

Să presupunem că, într-o perioadă de timp precizată, cererea dintr-un anumit tip de produs poate fi 0,1,2 sau 3 unităţi ştiindu-se că cereri mai mari de 3 unităţi de produs nu pot avea loc. Faptul că cererea va fi de 0, 1, 2 sau 3 unităţi de produs este influenţat de o serie de factori necontrolabili, sau necunoscuţi. Avîndu-se în vedere modul în care aceşti factori influenţeazăă nivelul cererii, se acceptă că cererea va fi de 0 unităţi de produs, (nu se va solicita acest produs), cu probabilitatea p_0 şi va fi de 1, 2, 3 unităţi de produs cu probabilitatea P_1, p_2, p_3, respectiv $(p_0+p_1+p_2+p_3=1)$. Acestei situaţii îi putem asocia un cîmp de probabilitate şi o variabilă aleatoare unidimensională definită peste acest cîmp de probabilitate, în felul următor:

Fie $\Omega=\{\omega_0,\omega_1,\omega_2,\omega_3\}$, $K=P(\Omega)$, şi probabilitatea $P:K\to\mathbb{R}$ definită prin: $P(\{\omega_i\})=p_i$, $i=0,1,2,3$, iar dacă $A\in K=P(\Omega)$ atunci $P(A)$ se determină după regulile obişnuite de calcul cu probabilităţi, de exemplu:

$$P(\{\omega_0,\omega_2\})=P(\{\omega_0\})+P(\{\omega_2\})=p_0+p_2$$

Definim acum funcţia $X:\Omega\to\mathbb{R}$ prin $X(\omega_i)=i$, $i=0,1,2,3$.

Deoarece $K = P(\Omega)$ rezultă că pentru oricare $a, a \in \mathbb{R}$, mulțimea $\{\omega / \omega \in \Omega, X(\omega) < a\} \in K$, deci X este o variabilă aleatoare peste cîmpul de probabilitate (Ω, K, P).

Remarcăm faptul că în cazul cînd se dă un cîmp de probabilitate (Ω, K, P) astfel încît $K = P(\Omega)$, atunci orice funcție $X: \Omega \to \mathbb{R}$ este o variabilă aleatoare unidimensională peste cîmpul de probabilitate dat.

Deşi în cele ce urmează vom lucra, aproape în exclusivitate, cu variabile aleatoare unidimensionale, dăm aici şi definiția unei variabile aleatoare multidimensionale.

Definiția 7.5.2.

Fie (Ω, K, P) un cîmp de probabilitate şi $X: \Omega \to \mathbb{R}^n$ o funcție definită pe Ω şi cu valori în spațiul euclidian n-dimensional. Dacă pentru orice $\omega \in \Omega$, notăm $X(\omega) = (X_1(\omega), \ldots, X_n(\omega))$, în care $X_i(\omega) \in \mathbb{R}$, $i = 1, \ldots, n$, atunci funcția X se spune că este o variabilă aleatoare n-dimensională, peste cîmpul de probabilitate (Ω, K, P), în cazul cînd pentru oricare n numere reale $a_1, \ldots, a_n$ are loc relația:

$$\{\omega \mid X_1(\omega) < a_1, \ldots, X_n(\omega) < a_n\} \in K \qquad 7.5.2$$

O prezentare succintă a variabilelor aleatoare multidimensionale va fi făcută în ultima parte a paragrafului. Revenind la variabile aleatoare unidimensionale se poate dovedi că, dacă (Ω, K, P) este un cîmp borelian de probabilitate, iar X este o variabilă aleatoare definită peste acest cîmp de probabilitate, sînt adevărate următoarele relații:

$$\{\omega \mid X(\omega) \le a\} \in K, \ (\forall) a \in \mathbb{R}$$
$$\{\omega \mid X(\omega) > a\} \in K, \ (\forall) a \in \mathbb{R} \qquad 7.5.3$$

Demonstrațiile acestor relații, cît şi demonstrația următoarei teoreme, pot fi găsite, de exemplu, în [9,10].

Teorema 7.5.1

Dacă X, Y sînt două variabile aleatoare, definite pe cîmpul de probabilitate (Ω, K, P), atunci funcțiile $c.X$, unde $c \in \mathbb{R}$, $|X|$, unde prin $|X|$ se

înţelege modulul funcţiei X, $\frac{X}{Y}$, dacă $Y(\omega) \neq 0$ pentru orice $\omega \in \Omega$, $X + Y$ şi $X \times Y$ sînt de asemeni variabile aleatoare peste cîmpul de probabilitate (Ω, K, P).

Vom defini o clasă specială de variabile aleatoare, numite *variabile aleatoare discrete.*

Definiţia 7.5.3.

O variabilă aleatoare X, definită peste un cîmp de probabilitate (Ω, K, P), se spune că este o variabilă discretă atunci cînd mulţimea valorilor sale, $X(\Omega)$, este o mulţime cel mult numărabilă, (mulţime discretă). Altfel spus, o variabilă aleatoare X definită peste cîmpul de probabilitate (Ω, K, P) se spune că este discretă dacă există o mulţime de formă $\{x_1, x_2, \ldots, x_n\} \subset \mathbb{R}$, sau $\{x_1, x_2, \ldots, x_n, \ldots\} \subset \mathbb{R}$ astfel încît $X(\Omega) = \{x_1, x_2, \ldots, x_n\}$, respectiv $X(\Omega) = \{x_1, x_2, \ldots, x_n, \ldots\}$. Atunci cînd mulţimea valorilor unei variabile aleatoare discrete este finită, variabila aleatoare respectivă se mai spune că este o variabilă aleatoare simplă. Ca exemplu de astfel de variabile aleatoare putem considera variabilele aleatoare construite în exemplele 7.5.1 şi 7.5.2.

Variabile aleatoare independente

În practică, pentru studierea proprietăţilor unei anumite caracteristici numerice a elementelor unei colectivităţi este de multe ori convenabil ca acestei caracteristici să i se pună în corespondenţă o variabilă aleatoare. Dacă la elementele unei astfel de colectivităţi ne interesează cel puţin două caracteristici numerice, desigur că este posibil ca între valorile acestor caracteristici să existe o relaţie de interdependenţă, sau o astfel de relaţie să nu existe. În cazul cînd între valorile caracteristicilor analizate nu există nici o relaţie de interdependenţă, vom spune, în mod natural, că aceste caracteristici sînt independente, iar în caz contrar vom spune că aceste caracteristici sînt dependente. Variabilele aleatoare, corespunzătoare acestor caracteristici ale elementelor colectivităţii, vom spune că sînt dependente sau independente după cum caracteristicile corespunzătoare ale colectivităţii sînt dependente, sau independente.

Pentru exemplificare, să considerăm că la membrii unei colectivităţi umane ne interesează vîrsta şi înălţimea, acestea fiind două caracteristici numerice ale membrilor colectivităţii. Fără îndoială că aceste două caracteristici nu pot fi considerate independente, fiind clar cel puţin faptul că, măcar pînă la o anumită vîrstă, înălţimea membrilor colectivităţii creşte odată cu vîrsta. Pe de altă parte, este clar şi faptul că nu toţi membrii colectivităţii considerate, care au o vîrstă dată, au şi aceeaşi înălţime sau invers, evoluţia acestor caracteristici ale membrilor colectivităţii fiind influenţată de diferiţi factori , cunoscuţi sau necunoscuţi. Putem interpreta aceste caracteristici ca nişte variabile aleatoare şi, deoarece între valorile acestor caracteristici există o relaţie de interdependenţă, vom spune că variabilelel aleatoare corespunzătoare sînt dependente. Acest mod de a defini interdependenţa a două variabile aleatoare este nesatisfăcător chiar şi în unele situaţii foarte simple. În scopul ilustrării acestui fapt ne vom referi la un exemplu aparent curios.

Să considerăm drept primă caracteristică numerică pentru un an calendaristic, recolta de seminţe de trifoi dintr-o anumită regiune, a doua caracteristică numerică fiind numărul de pisici din regiunea respectivă. Deşi, la prima vedere, între aceste două caracteristici nu pare a exista vreo legătură, biologii au stabilit o astfel de legătură şi încă destul de puternică. Această legătură constă în următoarele: s-a stabilit că polenizarea florilor de trifoi este asigurată în cea mai mare măsură de către bondari, care sînt în mumăr mic în anii în care există mulţi şoareci de cîmp, deoarece aceştia îi împuţinează; la rîndul lor şoarecii de cîmp sînt în număr mic atunci cînd sînt multe pisici, ceea ce se poate explica fără prea mari dificultăţi; interdependenţa dintre recolta de semninţe trifoi şi numărul de pisici dintr-un anumit an fiind astfel evidentă.

Vom da acum o definiţie precisă a independenţei a două variabile aleatoare, definite peste un acelaşi cîmp de probabilitate.

Definiţia 7.5.4.

Fie X, Y două variabile aleatoare definite peste cîmpul de probabilitate (Ω, K, P). Vom spune că aceste două variabile aleatoare sînt independente dacă pentru oricare $x, y \in \mathbf{R}$ are loc relaţia:

$$P(\{\omega \mid X(\omega) < x,\ Y(\omega) < y\}) = P(\{\omega \mid X(\omega) < x\})\ P(\{\omega \mid Y(\omega) < y\}) \qquad 7.5.4$$

Să observăm că variabilele aleatoare X, Y sînt independente dacă şi

numai dacă, pentru oricare două numere reale x, y, evenimentele $\{\omega \mid X(\omega) < x\}$, $\{\omega \mid Y(\omega) < y\}$, din cîmpul de probabilitate (Ω, K, P), sînt independente. În cazul cînd două variabile aleatoare nu sînt independente, se mai spune că sînt dependennte. Pentru simplificarea notaţiei să convenim ca de acum încolo, dacă X este o variabilă aleatoare peste un cîmp de probabilitate (Ω, K, P), să notăm $P(\{\omega \mid X(\omega) < x\})$ prin $P(X < x)$, (citim: „probabilitatea ca variabila aleatoare X să ia valori mai mici decît x"). Cu această convenţie putem spune că variabilele aleatoare X, Y, definite peste cîmpul de probabilitate (Ω, K, P), sînt independente dacă pentru orice $x, y \in \mathbb{R}$ are loc relaţia:

$$P(X < x,\ Y < y) = P(X < y) \cdot P(Y < y)$$

Exemplul 7.5.3.

Să presupunem că într-o urnă sînt 2 bile de culoare albă şi o bilă de culoare neagră şi din această urnă se extrag succesiv două bile, conform schemei bilei revenite, ceea ce ne interesează fiind culoarea bilelor extrase. Asociem acestei experienţe un cîmp finit de evenimente, (Ω, K), în care $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4\}$, $K = P(\Omega)$, evenimentele elementare $\{\omega_1\}, \ldots, \{\omega_4\}$ corespund respectiv situaţiilor în care prima bilă extrasă este albă şi a doua neagră, prima bilă neagră şi a doua albă, ambele bile de culoare neagră. Definind probabilitatea unui eveniment din acest cîmp de evenimente conform definiţiei clasice a probabilităţii se constată imediat că $P(\{\omega_1\}) = P(\{\omega_2\}) = \frac{2}{9}$, $P(\{\omega_3\}) = \frac{4}{9}$, $P(\{\omega_4\}) = \frac{1}{9}$, pentru determinarea probabilităţii celorlalte evenimente din acest cîmp de evenimente folosindu-se proprietăţile cunoscute ale probabilităţii (de exemplu $P(\{\omega_1, \omega_4\}) = P(\{\omega_1\}) + P(\{\omega_4\}) = \frac{1}{4}$).

Peste cîmpul de probabilitate (Ω, K, P) astfel definit să considerăm două variabile aleatoare X, Y, definite în felul următor:

$$X(\omega_1) = X(\omega_3) = 1 \quad , \quad X(\omega_2) = X(\omega_4) = 0$$

$$Y(\omega_2) = Y(\omega_3) = 1 \quad , \quad Y(\omega_1) = Y(\omega_4) = 0$$

Cu alte cuvinte variabila aleatoare X ia valoarea 1 atunci cînd prima bilă extrasă este de culoare albă şi ia valoarea zero atunci cînd prima bilă extrasă este de culoare neagră, iar variabila aleatoare Y ia valoarea 1 atunci cînd a doua bilă extrasă este de culoare albă şi ia valoarea zero atunci cînd a doua bilă extrasă este de culoare neagră.

Se constată imediat că putem scrie:

$$P(X<x,\ Y<y)=\begin{cases} P(\Phi), & \text{dacă } x\le 0 \text{ sau } y\le 0,\\ P(\{\omega_4\}), & \text{dacă } 0<x\le 1 \text{ şi } 0<y\le 1,\\ P(\{\omega_2,\omega_4\}), & \text{dacă } 0<x\le 1 \text{ şi } x>1,\\ P(\{\omega_1,\omega_4\}), & \text{dacă } 0<y\le 1 \text{ şi } x>1,\\ P(\{\omega_1,\omega_2,\omega_3,\omega_4\}), & \text{dacă } x>1 \text{ şi } y>1, \end{cases}$$

iar pe de altă parte:

$$P(X<x)=\begin{cases} P(\Phi);x\le 0\\ P(\{\omega_2,\omega_4\});0<x\le 1\\ P(\Omega);x>1 \end{cases}$$

$$P(Y<y)=\begin{cases} P(\Phi);y\le 0\\ P(\{\omega_1,\omega_4\});0<y\le 1\\ P(\Omega);y>1 \end{cases}$$

Se poate verifica imediat că pentru orice $x,y\in\mathbb{R}$ are loc relaţia $P(X<x,Y<y)=P(X<x)P(Y<y)$; de exemplu pentru $0<x\le 1$ şi $0<y\le 1$, $P(X<x,Y<y)=P(\{\omega_4\})=\frac{1}{9}$ iar $P(X<x)=P(\{\omega_2,\omega_4\})=\frac{2}{9}+\frac{1}{9}=\frac{1}{3}$, $P(Y<y)=P(\{\omega_1,\omega_4\})=\frac{2}{9}+\frac{1}{9}=\frac{1}{3}$. Deci variabilele aleatoare X şi Y sînt independente. Dacă peste acelaşi cîmp de evenimente definim o altă probabili-

tate P_1 astfel încît $P_1(\{\omega_1\}) = P_1(\{\omega_2\}) = P(\{\omega_3\}) = \frac{1}{3}$, $P_1(\{\omega_4\}) = 0$ şi considerăm aceleaşi variabile aleatoare X, Y definite ca mai sus, dar peste cîmpul de probabilitate (Ω, K, P_1) constatăm că şi de această dată nu sînt independente. Într-adevăr, dacă se consideră $0 < x \le 1$, $0 < y \le 1$ atunci $P_1(X < x, Y < y) = P_1(\omega_4) = 0$, iar $P_1(X < x) = \frac{1}{3}$, $P_1(Y < y) = \frac{1}{3}$ şi deci, măcar în aceste situaţii $P_1(X < x, Y < y) \neq P_1(X < x) P_1(Y < y)$.

În calitate de exerciţiu, sugerăm să se verifice faptul că probabilităţile evenimentelor $\{\omega_1\}$, $\{\omega_2\}$, $\{\omega_3\}$ sînt egale cu $\frac{1}{3}$ iar probabilităţile evenimentului $\{\omega_4\}$ este zero în cazul cînd cele două extrageri din urnă se fac conform schemei bilei nerevenite.

Definiţia 7.5.5.

Fie $X_1, X_2, \ldots, X_n$, n variabile aleatoare definite peste cîmpul de probabilitate (Ω, K, P), $n \ge 2$. Vom spune că aceste n variabile aleatoare sînt independente dacă pentru oricare $x_1, x_2, \ldots, x_n \in \mathbb{R}$ are loc relaţia:

$$P(X_1 < x_1, X_2 < x_2, \ldots, X_n < x_n) = \\ = P(X_1 < x_1) P(X_2 < x_2) \ldots P(X_n < x_n) \qquad 7.5.5$$

Observaţia 7.8. Remarcăm că dacă X, Y sînt două variabile aleatoare independente, definite peste un cîmp de probabilitate (Ω, K, P), iar $\alpha, \beta \in \mathbb{R}$ atunci $X + \alpha$, $Y + \beta$ sînt deasemenea variabile aleatoare independente şi, chiar mai mult, dacă $f: \mathbb{R} \to \mathbb{R}$, $g: \mathbb{R} \to \mathbb{R}$ sînt două funcţii reale de variabilă reală încît $f(X), g(Y)$ sînt variabile aleatoare peste cîmpul de probabilitate (Ω, K, P) atunci variabilele aleatoare $f(X), g(Y)$ sînt de asemeni independente (subliniem faptul că funcţiile $f(X): \Omega \to \mathbb{R}$, $g(Y): \Omega \to \mathbb{R}$ se definesc prin $f(X)(\omega) = f(X(\omega))$ şi $g(Y)(\omega) = g(Y(\omega))$).

Funcţia de repartiţie a unei variabile aleatoare

Să presupunem că cererea dintr-un anumit tip de produs poate fi considerată ca o variabilă peste un anumit cîmp de probabilitate. Este fără nici o îndoială util să se cunoască o măsură a şanselor ca cererea să se afle într-o anumită submulţime din valorile posibile. De exemplu dacă cererea poate lua valori întregi cuprinse între 0 şi 100, ar putea fi util să ştim ce şanse sînt ca cererea să depăşească 50 de unităţi de produs, sau să nu depăşească 25 unităţi de produs, eventual ar putea să ne intereseze probabilitatea ca cererea să se afle în oricare dintre situaţiile posibile. Interpretînd cererea ca o variabilă aleatoare, răspunsul la astfel de întrebări poate fi obţinut folosind *funcţia de repartiţie* a variabilei aleatoare corespunzătoare. Vom defini acum funcţia de repatiţie pentru o variabilă aleatoare oarecare (unidimensională).

Definiţia 7.5.6.

Fie X o variabilă aleatoare definită peste cîmpul de probabilitate (Ω, K, P). Prin funcţia de repartiţie a variabilei aleatoare X se înţelege funcţia $F: \mathbb{R} \to [0,1]$, definită prin:

$$F(x) = P(X < x), \ (\forall) x \in \mathbb{R} \qquad 7.5.6$$

Se poate dovedi că funcţia de repartiţie F a unei variabile aleatoare X, definită peste un cîmp borelian de probabilitate (Ω, K, P), are următoarele proprietăţi:

1. Dacă $x_1, x_2 \in \mathbb{R}$, $x_1 < x_2$ atunci $F(x_1) \leq F(x_2)$, adică funcţia de repartiţie este o funcţie monoton nedescrescătoare.

2. $\lim\limits_{x \nearrow x_0} F(x) = F(x_0)$, $(\forall) x_0 \in \mathbb{R}$, adică funcţia de repartiţie a unei variabile aleatoare este o funcţie continuă la stînga în orice punct.

3. $\lim\limits_{x \to +\infty} F(x) = 1$ şi $\lim\limits_{x \to -\infty} F(x) = 0$.

4. Dacă $x_1, x_2 \in \mathbb{R}$ astfel încît $x_1 < x_2$ atunci:

$$P(x_1 \leq X < x_2) = F(x_2) - F(x_1)$$

5. Dacă $a \leq \mathbb{R}$ atunci $P(X = a) = \lim\limits_{x \searrow a} F(x) - F(a)$.

De asemeni, se poate arăta că, dacă $F:\mathbb{R}\to[0,1]$ este o funcţie satisfăcînd proprietăţile 1-3 de mai sus atunci există un cîmp de probabilitate $(\Omega, K, ,P)$ şi o variabilă aleatoare X peste acest cîmp de probabilitate, astfel încît F să fie funcţia de repartiţie a variabilei aleatoare X.

Vom dovedi aici doar proprietăţile 1 şi 4, justificarea celorlalte proprietăţi se poate găsi de exemplu în [10].

Demonstraţia proprietăţii 1: Fie $x_1, x_2 \in \mathbb{R}$ astfel încît $x_1 < x_2$. Deoarece $F(x_1) = P(X<x)$, $F(x_2) = P(X<x_2)$, iar $P(X<x_2) = P(X<x_1)$, sau $x_1 \le X \le x_2$ din incompatibilitatea evenimentelor $\{\omega \mid X(\omega) < x_1\}, \{\omega \mid x_1 \le X(\omega) < x_2\}$ urmează $F(x_2) = P(X<x_1) + P(x_1 \le X < x_2)$. Dar $P(X<x_1) = F(x_1)$ iar $\mathcal{P}(x_1 \le X < x_2) \ge 0$, ceea ce arată că $F(x_2) \ge F(x_1)$ şi cu aceasta prima proprietate este justificată.

Demonstraţia proprietăţii 4: Fie $x_1, x_2 \in \mathbb{R}$, $x_1 < x_2$. Urmînd acelaşi drum ca şi în demonstraţia proprietăţii 1, se constată că:

$$P(X<x_2) = P(X<x_1) + P(x_1 X < x_2)$$

şi ţinînd cont de definiţia funcţiei de repartiţie a variabilei aleatoare X, urmează:

$$P(x_1 \le X < x_2) = F(x_2) - F(x_1)$$

ceea ce justifică şi această proprietate a funcţiei de repartiţie a unei variabile aleatoare.

Exemplul 7.5.4.

Să considerăm că cererea pentru un anumit tip de piese de schimb în decurs de un trimestru poate fi 0, 1 sau 2 unităţi convenţionale, probabilităţile ca cererea să fie de 0, 1, 2 unităţi fiind 0,1; 0,5; 0,4 respectiv. În această situaţie putem interpreta cererea ca o variabilă aleatoare, X, peste cîmpul de probabilitate (Ω, K, P), în care $\Omega = \{\omega_0, \omega_1, \omega_2\}$, $K = P(\Omega)$, $P(\{\omega_0\}) = 0,1$, $P(\{\omega_1\}) = 0,5$, $P(\{\omega_2\}) = 0,4$, iar $X(\omega_0) = 0$, $X(\omega_1) = 1$, $X(\omega_2) = 2$.

Notînd cu F funcţia de repartiţie a variabilei aleatoare X atunci, conform definiţiei 7.56, urmează:

$$F(x) = \begin{cases} 0 & ; \ x \le 0 \\ 0,1 & ; \ 0 < x \le 1 \\ 0,6 & ; \ 1 < x \le 2 \\ 1 & ; \ x > 2 \end{cases}$$

De exemplu, dacă $1 < x \le 2$ atunci:

$$F(x) = P(X < x) = P(\{\omega_0, \omega_1\}) = 0,1 + 0,5 = 0,6$$

Dacă dorim să găsim probabilitatea ca cererea din acest tip de piese de schimb să fie de cel puţin o unitate convenţională, în decurs de un trimestru, atunci aceasta este:

$$P(X \ge 1) = 1 - P(X < 1) = 1 - F(1) = 0,9$$

Funcţia de repartiţie a unei variabile aleatoare discrete

În capitolul precedent am definit o variabilă aleatoare discretă ca fiind o variabilă aleatoare pentru care mulţimea valorilor sale este o mulţime discretă.

Pentru variabilele aleatoare discrete funcţia de repartiţie are o formă cu totul particulară, formă caracteristică pentru acest tip de variabilele aleatoare. În continuare vom construi funcţia de repatiţie a unei variabile aleatoare simple, pentru o variabilă aleatoare discretă cu o mulţime numărabilă de valori construcţia fiind cu totul asemănătoare.

Fie X o variabilă aleatoare simplă definită peste cîmpul de probabilitate (Ω, K, P) şi fie $X(\Omega) = \{x_1, x_2, ..., x_n\}$. Fără a se micşora generalitatea, putem presupune că numerele x_i, $i = 1, 2, ..., n$, sînt ordonate crescător, adică $x_1 < x_2 < ... < x_n$. Să notăm prin A_i, $1 \le i \le n$, evenimentul definit prin relaţia:

$$A_i = \{\omega / \omega \in \Omega \ , X(\omega) = x_i\}$$

Este atunci clar că $A_i \neq \Phi$, $i = 1, .., n$, $\bigcup_{i=1}^{n} A_i = \Omega$ şi, în plus, $A_i \cap A_j = \Phi$, pentru oricare i, j astfel încît $i \neq j$. Să mai notăm cu p_i, $1 \le i \le n$, probabilitatea evenimentului A_i, adică $p_i = P(A_i)$, $i = 1, ..., n$.

Considerînd acum un număr real oarecare x, constatăm că valoarea corespunzătoare, $F(x)$, a funcţiei de repartiţie a variabilei aleatoare X este dată de relaţia:

$$F(x)=\begin{cases} 0 & ;\ x\le x_1 \\ p_1 & ;\ x_1<x\le x_2 \\ \vdots & \\ \sum_{i=1}^{k-1} p_i & ;\ x_{k-1}<x\le x_k \\ \vdots & \\ \sum_{i=1}^{n-1} p_i & ;\ x_{n-1}<x\le x_n \\ \sum_{i=1}^{n} p_i=1 & ;\ x>x_n \end{cases} \qquad 7.5.7$$

Se poate deci spune că funcţia de repartiţie, sau mai pe scurt *repartiţia* unei variabile de tip discret este complet determinată dacă se cunosc valorile x_i, $i=1,2,\ldots$, ale acestei variabile aleatoare şi probabilităţile evenimentelor A_i, $P(A_i)=p_i$, $i=1,2,\ldots$, definite ca mai sus. Vom mai spune, în acest caz, că p_i este probabilitatea cu care variabile aleatoare X ia valoarea x_i. Prin urmare, repartiţia unei variabile aleatoare simple, X, poate fi dată printr-un tabel de forma:

$$\begin{pmatrix} x_1, & x_2, \ldots, & x_n \\ p_1, & p_2, \ldots, & p_n \end{pmatrix}$$

în care în prima linie se află valorile variabilei aleatoare X, iar în linia a doua se află probabilitatea corespunzătoare, cu care variabila aleatoare ia aceste valori. Un astfel de tabel se numeşte *tablou de repartiţie* al variabilei aleatoare considerate. În cazul cînd variabila aleatoare X este discretă, dar nu este simplă, tabloul de repartiţie asociat ei va fi de forma:

$$\begin{pmatrix} x_1, & x_2, \ldots, & x_n, \ldots \\ p_1, & p_2, \ldots & p_n, \ldots \end{pmatrix}$$

BCU "LUCIAN BLAGA" UNIVERSITY LIBRARY

Mai notăm faptul că un tabel de forma:

$$\begin{pmatrix} x_1, x_2, \dots \\ p_1, p_2, \dots \end{pmatrix}$$

poate fi tabloul de reapartiţie al unei variabile aleatoare discrete dacă sînt îndeplinite condiţiile:

a) $x_i, p_i \in \mathbb{R}, i = 1, 2, \dots$

b) $p_i \geq 0, i = 1, 2 \dots$

c) $\sum_i p_i = 1$

cu observaţia că, în cazul cînd numărul valorilor p_i nu este finit, prin $\sum_i p_i = 1$ înţelegem că seria $\sum_{i=1}^{\infty} p_i$ este convergentă la 1.

Exemplul 7.5.5.

Dacă se consideră tabelul:

$$\begin{pmatrix} 0 & 1 & 2 \dots\dots n \dots\dots \\ \frac{1}{2} & \frac{1}{2^2} & \frac{2}{2^3} \dots \frac{1}{2^{n-1}} \dots \end{pmatrix},$$

atunci acest tabel poate fi considerat drept tablou de repartiţie pentru o variabilă aleatoare discretă, deoarece numerele din linia a doua sînt toate nenegative, iar suma lor este:

$$\sum_{n=1}^{\infty} \frac{1}{2^n} = \lim_{n \to \infty} \frac{\frac{1}{2} - \frac{1}{2^{n+1}}}{1 - \frac{1}{2}} = 1.$$

Dacă însă se consideră tabelul:

$$\begin{pmatrix} 0 & 1 & 2 \dots\dots n \dots\dots \\ \frac{1}{3} & \frac{1}{3^2} & \frac{1}{3^3} \dots \frac{1}{3^{n-1}} \dots \end{pmatrix}$$

acesta nu poate constitui tabloul de repartiţie pentru o variabilă aleatoare discretă deoarece, deşi toate numerele din linia a doua sînt negative, suma lor

este:

$$\sum_{n=1}^{\infty} \frac{1}{3^n} = \lim_{n\to\infty} \frac{\frac{1}{3} - \frac{1}{3^{n+1}}}{1 - \frac{1}{3}} = \frac{1}{2} \neq 1$$

În continuare, tabloul de repatiţie al unei variabile aleatoare discrete îl vom scrie de obicei sub forma

$$\begin{pmatrix} x_i \\ p_i \end{pmatrix}, \quad i = 1, 2, \ldots, n$$

în cazul variabilelor aleatoare simple, şi sub forma:

$$\begin{pmatrix} x_i \\ p_i \end{pmatrix}, \quad i = 1, 2, \ldots$$

în cazul variabilelor aleatoare discrete cu o mulţime numărabilă de valori.

Variabile aleatoare de tip continuu. Funcţie de densitate

Pe lîngă variabilele aleatoare de tip discret, în practică este deseori necesară utilizarea a încă o categorie de variabile aleatoare, numite variabile aleatoare de tip continuu, sau variabile aleatoare cu *repartiţie de tip continuu*. În mare vorbind, se spune că o variabilă aleatoare este de tip continuu într-un interval (a, b), dacă poate lua orice valoare din acest interval. O definiţie mai exactă a unei variabile aleatoare de tip continuu o dăm în cele ce urmează.

Definiţia 7.5.7.

Fie X o variabilă aleatoare definită peste un cîmp de probabilitate (Ω, K, P). Vom spune că variabila aleatoare X este o variabilă aleatoare de tip continuu dacă există o funcţie f, $f: \mathbb{R} \to \mathbb{R}$, astfel încît:

$$F(x) = \int_{-\infty}^{x} f(t)\,dt, \quad (\forall)\, x \in \mathbb{R}$$

unde F este funcţia de repartiţie a variabilei aleatoare X.

Definiţia 7.5.8.

Dacă X este o variabilă aleatoare de tip continuu şi dacă funcţia sa de repartiţie, F, este dată prin:

$$F(x) = \int_{-\infty}^{x} f(t)\,dt,\ (\forall)x \in \mathbb{R}$$

iar $f(x) \geq 0,\ (\forall)x \in \mathbb{R}$, atunci funcţia f, $f: \mathbb{R} \to \mathbb{R}$, se numeşte funcţie de densitate de probabilitate pentru variabila aleatoare X, sau pe scurt funcţie de densitate a variabilei aleatoare X.

Pe baza celor două definiţii de mai sus, se poate arăta că o variabilă aleatoare este de tip continuu dacă ea are funcţie de densitate determină în mod unic repartiţia aleatoare considerate.

Din definiţia funcţiei de densitate pentru o variabilă aleatoare de tip continuu urmează că orice funcţie de densitate, f, satisface condiţiile:

1. $f(x) \geq 0,\ (\forall)x \in \mathbb{R}$ 7.5.8

2. $\int_{-\infty}^{+\infty} f(x)dx = 1$

Pe de altă parte, se poate arăta că, dacă se dă o funcţie $f: \mathbb{R} \to \mathbb{R}$ care îndeplineşte condiţiile (7.5.8) atunci funcţia $F: \mathbb{R} \to [0,1]$, definită prin:

$$F(x) = \int_{-\infty}^{x} f(t)\,dt,\ (\forall)x \in \mathbb{R} \qquad 7.5.9$$

este funcţia de repartiţie pentru o anumită variabilă aleatoare, definită peste un cîmp de probabilitate convenabil ales. Remarcăm faptul că în cazul în care funcţia de repartiţie, F, a unei variabile aleatoare X este derivabilă atunci variabila aleatoare X este o variabilă aleatoare de tip continuu pentru care derivata funcţiei F este funcţie de densitate a variabilei aleatoare X.

Teorema 7.5.2.

Dacă X este o variabilă aleatoare de tip continuu cu funcţia de densitate f, atunci variabilele aleatoare $cX, X^2, X+c, (c \in \mathbb{R}, c \neq 0)$, sînt deasemeni variabile aleatoare de tip continuu, funcţiile lor de densitate $\varphi_1, \varphi_2, \varphi_3$, respectiv, fiind definite prin:

$$\varphi_1(x) = \frac{1}{|c|} f(\frac{x}{c}),\ (\forall) x \in \mathbb{R}$$

$$\varphi_2(x) = \begin{cases} \frac{1}{2\sqrt{x}} f(\sqrt{x}) + f(-\sqrt{x}), & x > 0 \\ 0, & x \le 0 \end{cases} \qquad 7.5.10$$

$$\varphi_3(x) = f(x-c),\ (\forall) x \in \mathbb{R}$$

$$\varphi_3(x) = f(x-c),\ (\forall) x \in \mathbb{R} \qquad 7.5.11$$

unde $|c|$ este modulul numărului c.

Demonstraţie: Vom demonstra numai faptul că variabila aleatoare $cX, (c \in \mathbb{R},\ c \neq 0)$, este o variabilă aleatoare de tip continuu, cu funcţia de densitate φ_1. Faptul că φ_2, φ_3 sînt funcţii de densitate pentru variabilele aleatoare X^2, $X + c$, respectiv, se dovedeşte în mod asemănător, motiv pentru care lăsăm ca exerciţiu justificarea completă a teoremei.

Să dovedim deci că funcţia φ_1 este funcţie de densitate a variabilei aleatoare cX.

Conform definiţiei 7.5.2, funcţia de repartiţie a variabilei aleatoare cX, $F_1 : \mathbb{R} \to [0,1]$, se defineşte prin:

$$F_1(x) = P(cX < x),\ (\forall) x \in \mathbb{R}$$

Deci, pentru a dovedi că φ_1 este funcţie de densitate a variabilei aleatoare cX este suficient să arătăm că are loc relaţia

$$P(cX < x) = \int_{-\infty}^{x} \frac{1}{|c|} f(\frac{t}{c})\, dt \qquad 7.5.12$$

deoarece $\varphi_1(x) \ge 0,\ (\forall) x \in \mathbb{R}$, rezultă imediat din faptul că $|c| > 0$ şi $f(\frac{x}{c}) \ge 0,\ (\forall) x \in \mathbb{R}$. Dar, este clar că au loc relaţiile:

$$P(cX < x) = P(X < \frac{x}{c}),\ (\forall) x \in \mathbb{R} \quad \textit{dacă} \quad c > 0$$

$$P(cX<x)=P(X>\frac{x}{c}),\ (\forall)x\in\mathbb{R},\ \ dacă\ \ c<0$$

adică putem scrie:

$$P(cX<x)=\begin{cases}F(\frac{x}{c}), & dacă\ \ c>0\\ 1-F(\frac{x}{c}), & dacă\ \ c<0\end{cases} \qquad 7.5.13$$

unde F este funcţia de repartiţie a variabilei aleatoare X. Pe de altă parte, deoarece f este funcţia de densitate pentru variabila aleatoare X, rezultă

$$F\left(\frac{x}{c}\right)=\int_{-\infty}^{\frac{x}{c}} f(t)\,dt,\ (\forall)x\in\mathbb{R},\ (\forall)c\neq 0 \qquad 7.5.14$$

Să presupunem acum $c>0$. În acestă situaţie $|c|=c$ şi, în urma substituţiei $u=ct$ în integrala din membrul doi al egalităţii (7.5.14), se obţine:

$$P(cX<x)=F(\frac{x}{c})=\int_{-\infty}^{x}\frac{1}{c}f(\frac{u}{c})\,du$$

adică putem spune că relaţia (7.5.14) are loc în cazul $c>0$.

Dacă $c<0$ atunci $|c|=-c$ şi făcînd din nou substituţia $u=ct$, dar de această dată în integrala din egalitatea:

$$1-F(\frac{x}{c})=\int_{\frac{x}{c}}^{+\infty} f(t)\,dt$$

obţinem, ţinînd cont şi de (7.5.13),

$$P(cX<x)=\int_{x}^{-\infty}\frac{1}{c}f(\frac{u}{c})\,du=-\int_{-\infty}^{x}\frac{1}{c}f(\frac{u}{c})=\int_{-\infty}^{x}\frac{1}{|c|}f(\frac{u}{c})du$$

În concluzie, relaţia (7.5.12) are loc atît în cazul $c>0$ cît şi în cazul $c<0$; deci φ_1 este funcţie de densitatea pentru variabila aleatoare cX.

BCU IAŞI CENTRAL UNIVERSITY LIBRARY

Variabile aleatoare n-dimensionale

În cele precedente am observat că în unele situaţii este util să asociem elementelor mulţimii Ω, dintr-un cîmp de probabilitate (Ω, K, P), ansambluri ordonate de numere reale. După cum se ştie, o funcţie $X: \Omega \to \mathbb{R}^n$ asociază fiecărui element din Ω un ansamblu ordonat de n numere reale, funcţia X putînd fi scrisă şi sub forma $X = (X_1,...,X_n)$, în care X_i, $i = 1,...,n$, sînt funcţii definite pe Ω şi cu valori reale, $X_i: \Omega \to \mathbb{R}$, $i = 1,...,n$. Este atunci naturală următoarea definiţie pentru o variabilă aleatoare n-dimensională (a se vedea definiţia 7.5.2).

Definiţia 7.5.9.

Fie (Ω, K, P) un cîmp de probabilitate şi X o funcţie definită pe Ω şi cu valori în $\mathbb{R}^n$, $X: \Omega \to \mathbb{R}^n$. Vom spune că funcţia X este o variabilă aleatoare n-dimensională peste cîmpul de probabilitate (Ω, K, P) dacă pentru oricare n numere reale a_i, $i = 1,...,n$, are loc relaţia:

$$\{\omega \mid \omega \in \Omega,\ X_i(\omega) < a_i,\ i = 1,...,n\} \in K \qquad 7.5.15$$

Este clar că atunci cînd X_i, $i = 1,...,n$, sînt variabile aleatoare unidimensionale definite peste un acelaşi cîmp de probabilitate (Ω, K, P), urmează că $X = (X_1,...,X_n)$ este o variabilă aleatoare n-dimensională definită peste (Ω, K, P). Reciproc, dacă $X = (X_1,...,X_n)$ este o variabilă aleatoare n-dimensională peste un cîmp de probabilitate (Ω, K, P) atunci funcţiile $X_i: \Omega \to \mathbb{R}$, $i = 1, ..., n$ sînt variabile aleatoare unidimensionale peste (Ω, K, P), fapt care rezultă din relaţiile:

$$\{\omega \mid X_i(\omega) < x\} = \bigcup_{k=1}^{\infty} \{\omega \mid X_i(\omega) < x,\ X_j(\omega) < k,\ (\forall) j \neq i\} \qquad 7.5.16$$

şi din faptul că $X = (X_1,...,X_n)$ este variabilă aleatoare n-dimensională peste (Ω, K, P).

Definiţia 7.5.10.

Se spune că o variabilă aleatoare n-dimensională X, definită peste un cîmp de probabilitate (Ω, K, P), este o variabilă aleatoare de tip discret dacă mulţimea $X(\Omega)$ este o mulţime discretă în $\mathbf{R}^n$.

Vom defini acum funcţia de repartiţie a unei variabile aleatoare n-dimensionale.

Definiţia 7.5.11.

Dacă $X = (X_1, ..., X_n)$ este o variabilă aleatoare n-dimensională peste cîmpul de probabilitate (Ω, K, P) atunci prin funcţia de repartiţie a acestei variabile aleatoare înţelegem funcţia $F : \mathbf{R}^n \to \mathbf{R}$, definită prin:

$$F(x_1, ..., x_n) = P(X_1 < x_1, ..., X_n < x_n), \ (\forall)(x_1, ..., x_n) \in \mathbf{R}^n \qquad 7.5.17$$

unde prin $P(X_1 < x_1, ..., X_n < x_n)$ se înţelege probabilitatea evenimentului $\{\omega \mid \omega \in \Omega, \ X_i(\omega) < x_i, \ i = 1, ..., n\} \in K$

Se spune că repartiţia unei variabile aleatoare n-dimensionale este cunoscută dacă se cunoaşte funcţia de repartiţie a variabilei aleatoare respective.

Ca şi în cazul variabilelor aleatoare unidimensionale de tip discret, pentru variabilele aleatoare n-dimensionale de tip discret repartiţia se poate printr-un tabel în care sînt precizate valorile variabilei aleatoare (vectori din $\mathbf{R}^n$) şi probabilităţile cu care variabila aleatoare ia aceste valori.

Definiţia 7.5.12.

O variabilă aleatoare n-dimensională X_1, se spune că este de tip continuu dacă există o funcţie $f: \mathbf{R}^n \to \mathbf{R}^n$ astfel încît funcţia de repartiţie F, pentru variabila aleatoare dată să se scrie sub forma:

$$F(x_1, ..., x_n) = \int_{-\infty}^{x_1} ... \int_{-\infty}^{x_n} f(u_1, ..., u_n) \, du_1 ... du_n \qquad 7.5.18$$

$$(\forall)(x_1, ..., x_n) \in \mathbf{R}^n$$

În cele ce urmează ne vom ocupa numai de variabile aleatoare bidimensionale, adică de variabile aleatoare ale căror valori se află în spaţiul vectorial real cu două dimensiuni, $\mathbf{R}^2$.

Variabile aleatoare bidimensionale

Conform cu cele spuse mai înainte, o variabilă aleatoare bidimensională, definită peste un cîmp de probabilitate (Ω, K, P), este o funcţie $X = (X_1, X_2) : \Omega \to \mathbb{R}^2$ cu proprietatea că pentru oricare două numere reale x_1, x_2 are loc relaţia:

$$\{\omega \mid \omega \in \Omega,\ X_1(\omega) < x_1,\ X_2(\omega) < x_2\} \in K$$

Pe de altă parte, o variabilă aleatoare bidimensională se poate considera ca fiind o pereche ordonată de variabile aleatoare unidimensionale, motiv pentru care o variabilă aleatoare bidimensională se mai numeşte şi vector aleator bidimensional. Pentru a evita folosirea indicilor, vom nota în mod obişnuit o variabilă aleatoare bidimensională prin perechea ordonată (X, Y) în care X şi Y sînt considerate ca variabile aleatoare unidimensionale definite peste un acelaşi cîmp de probabilitate (Ω, K, P). Avînd în vedere definiţia 7.5.10 o variabilă aleatoare bidimensională se spune că este de tip discret dacă mulţimea valorilor sale este o mulţime discretă în $\mathbb{R}^n$. Cu alte cuvinte, o variabilă aleatoare bidimensională (X, Y) definită peste un cîmp de probabilitate (Ω, K, P), este de tip discret atunci cînd există două mulţimi de numere naturale $I, J, I \subset \mathbb{N}, J \subset \mathbb{N}$, astfel încît:

$$(X(\omega), Y(\omega)) \in \{(x_i, y_j) \mid i \in I, j \in J\},\ (\forall)\ \omega \in \Omega$$

unde mulţimile $\{x_i \mid i \in I\}$, $\{y_j \mid j \in J\}$ sînt mulţimi precizate de numere reale.

Dacă se notează $P(X = x_i,\ Y = y_j) = p_{ij}$, $i \in I, j \in I$, atunci repartiţia variabilei aleatoare bidimensionale de tip discret (X, Y) este complet precizată prin tabloul:

$$\begin{pmatrix} (x_i, y_j) \\ p_{ij} \end{pmatrix};\quad i \in I, j \in J$$

La fel ca în cazul unidimensional, se poate verifica fără dificultate că numerele $p_{ij}, i \in I, j \in J$, îndeplinesc condiţiile:

a) $p_{ij} \geq 0,\ i \in I,\ j \in J$

b) $\sum_{i \in I} \sum_{j \in J} p_{ij} = 1$

Dacă se are în vedere definiţia 7.5.12, urmează că o variabilă aleatoare bidimensională (X, Y) se spune că este de tip continuu dacă există o funcţie $f: \mathbb{R}^2 \to \mathbb{R}^2$ astfel încît funcţia de repartiţie, F, a acestei variabile aleatoare se poate exprima prin:

$$F(x,y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f(u,v)\, du\, dv, \ (\forall)\,(x,y) \in \mathbb{R}^2$$

funcţia f purtînd denumirea de funcţie de densitate pentru variabila aleatoare bidimensională (X, Y).

Pe baza observaţiilor făcute după ce am definit noţiunea de variabilă aleatoare n-dimensională, urmează că în cazul cînd X, Y sînt variabile aleatoare unidimensionale peste un cîmp de probabilitate (Ω, K, P) atunci perechea (X, Y) este o variabilă aleatoare bidimensională peste (Ω, K, P), iar dacă (X, Y) este o variabilă aleatoare bidimensională, peste (Ω, K, P) atunci X, Y sînt variabile aleatoare unidimensionale peste (Ω, K, P). Are sens atunci următoarea definiţie:

Definiţia 7.5.13.

Dacă (X, Y) este o variabilă aleatoare bidimensională peste un cîmp de probabilitate (Ω, K, P) atunci variabilele aleatoare unidimensionale X, Y se numesc variabile aleatoare *marginale* asociate variabilei aleatoare bidimensionale (X, Y), sau repartiţii marginale asociate repartiţiei bidimensionale (X, Y).

Prin următoarele două teoreme vom pune în evidenţă legătura între repartiţiile a două variabile aleatoare unidimensionale independente şi repartiţia vectorului aleator asociat acestora.

Teorema 7.5.3.

Fie X, Y două variabile aleatoare unidimensionale de tip discret, peste un cîmp de probabilitate (Ω, K, P), cu tablourile de repartiţie:

$$X: \begin{pmatrix} x_i \\ p_i \end{pmatrix}, i \in I \subset N \ ; \ Y: \begin{pmatrix} y_i \\ q_i \end{pmatrix}, j \in J \subset N$$

Atunci, dacă variabilele aleatoare X, Y sînt independente, urmează că variabila aleatoare bidimensională (X, Y) are tabloul de repartiţie

$$\begin{pmatrix} (x_i, y_i) \\ p_i q_i \end{pmatrix}; \; i \in I, j \in J$$

Demonstraţie: Este clar că mulţimea valorilor variabilei aleatoare bidimensionale (X, Y) este constituită din mulţimea perechilor (x_i, y_i), $i \in I, j \in J$, atunci cînd $X(\Omega) = \{x_i | i \in I\}$, $Y(\Omega) = \{y_i | j \in J\}$.

Pentru a completa demonstraţia rămîne să dovedim că are loc relaţia:

$$P(X = x_i, Y = y_i) = p_i q_j, \; i \in I, j \in J$$

relaţie adevărată deoarece variabilele aleatoare X, Y, prin ipoteză, sînt independente. Cu aceasta teorema este demonstrată.

Teorema 7.5.4.

Fie X, Y două variabile aleatoare unidimensionale de tip continuu, definite peste un cîmp de probabilitate (Ω, K, P). Atunci, în ipoteza că X şi Y sînt independente, dacă $f_1 : \mathbb{R} \to \mathbb{R}$, $f_2 : \mathbb{R} \to \mathbb{R}$ sînt funcţiile de densitate ale variabilelor aleatoare X, Y, respectiv, urmează că variabila aleatoare bidimensională (X, Y) are funcţie de densitate $f : \mathbb{R}^2 \to \mathbb{R}^2$ definită prin:

$$f(x,y) = f_1(x) f_2(y); \; (\forall)(x, y) \in \mathbb{R}^2$$

Demonstraţie: Fie $F : \mathbb{R}^2 \to \mathbb{R}$ funcţia de repartiţie a variabilei aleatoare bidimensionale (X, Y) şi $F_1 : \mathbb{R} \to \mathbb{R}$, $F_2 : \mathbb{R} \to \mathbb{R}$, funcţiile de repartiţie ale variabilelor aleatoare unidimensionale X, Y, respectiv. Avînd în vedere ipoteza teoremei, putem atunci scrie

a) $F_1(x) = \int_{-\infty}^{x} f_1(u)\, du, \; (\forall) x \in \mathbb{R}$

b) $F_2(y) = \int_{-\infty}^{y} f_2(v)\, dv, \; (\forall) y \in \mathbb{R}$

c) $F(x,y) = F_1(x) F_2(y), \; (\forall)(x,y) \in \mathbb{R}^2$

Pe baza acestor relaţii şi a proprietăţilor integralei duble rezultă că putem scrie:

$$F(x,y)=\int_{-\infty}^{x}\int_{-\infty}^{y} f_1(u)f_2(v)\,du\,dv,\quad (\forall)\,(x,y)\in\mathbb{R}^2$$

ceea ce dovedeşte că funcţia $f:\mathbb{R}^2\to\mathbb{R}^2$ definită prin $f(x,y)=f_1(x)f_2(y)$ este funcţie de densitate pentru variabila aleatoare bidimensională (X,Y) şi teorema este demonstrată.

Vom demonstra acum două teoreme prin care se stabileşte legătura în repartiţia unei variabile aleatoare bidimensionale de tip discret sau de tip continuu şi repartiţiile variabilelor aleatoare marginale asociate acesteia.

Teorema 7.5.5.

Fie (X,Y) o variabilă aleatoare bidimensională de tip discret, cu tabloul de repartiţie:

$$\begin{pmatrix}(x_i,y_i)\\ p_{ij}\end{pmatrix};\quad i\in I\subset\mathbb{N}, j\in J\subset\mathbb{N}$$

Atunci variabilele aleatoare marginale X, Y au tablourile de repartiţie:

$$X:\begin{pmatrix}x_i\\ p_i\end{pmatrix},\ i\in I\ ;\quad Y:\begin{pmatrix}y_i\\ q_i\end{pmatrix},\ j\in J$$

unde $p_i=\sum_{j\in J}p_{ij},\ (\forall)\,i\in I$ şi $q_i=\sum_{i\in I}p_{ij},\ (\forall)\,j\in J$.

Demonstraţie: Avînd în vedere definiţia variabilelor aleatoare marginale X,Y este clar că, în cazul cînd $(X,Y)(\Omega)=\{(x_i,y_i)\mid i\in I, j\in J\}$, au loc relaţiile:

$$X(\Omega)=\{x_i\mid i\in I\},\ Y(\Omega)=\{y_i\mid j\in J\}$$

şi deci , pentru a completa demonstraţia teoremei, rămîne de dovedit că $P(X=x_i)=p_i$, $(\forall)\,i\in I$ şi $P(Y=y_i)=q_i$, $(\forall)\,j\in J$, unde p_i,q_i, $i\in I$, $j\in J$ sînt definiţi în enunţul teoremei.

Deoarece $\{\omega\mid Y(\omega)\in\{y_i\mid j\in J\}\}=\Omega$ putem scrie:

$$\mathcal{P}(X=x_i)=\mathcal{P}(X=x_i, Y\in\{y_j\mid j\in J\}),\ (\forall)\,i\in I$$

şi pe baza faptului că evenimentele $\{\omega \mid Y(\omega) = y_j\}, j \in J$, sînt incompatibile două cîte două urmează:

$$P(X = x_i) = \sum_{j \in J} P(X = x_i,\ Y = y_i),\ (\forall) i \in I$$

sau echivalent:

$$P(X = x_i) = \sum_{j \in J} p_{ij},\ (\forall) i \in I$$

În mod cu totul asemănător se arată că are loc:

$$P(Y = y_i) = \sum_{i \in I} p_{ij},\ (\forall) j \in J$$

ceea ce completează demonstraţia teoremei.

Teorema 7.5.6.

Fie (X,Y) o variabilă aleatoare tridimensională de tip continuu cu funcţia de densitate $f: \mathbb{R}^2 \to \mathbb{R}^2$. Atunci variabilele aleatoare marginale X, Y au funcţiile de densitate f_1, f_2, respectiv, definite prin:

$$f_1(x) = \int_{-\infty}^{+\infty} f(x,y)\,dy \quad , \quad (\forall) x \in \mathbb{R}$$

$$f_2(y) = \int_{-\infty}^{+\infty} f(x,y)\,dx \quad , \quad (\forall) y \in \mathbb{R}$$

Demonstraţie: Deoarece $\{\omega \mid Y(\omega) < +\infty\} = \Omega$ şi $\{\omega \mid X(\omega) < +\infty\} = \Omega$ rezultă că putem scrie:

$$P(X < x) = P(X < x,\ Y < +\infty),\ (\forall) x \in \mathbb{R}$$

$$P(Y < y) = P(X < +\infty, Y < y),\ (\forall) y \in \mathbb{R}$$

Dacă notăm prin F, F_1, F_2 funcţiile de repartiţie ale variabilelor aleatoare (X,Y), X, Y respectiv din relaţiile de mai sus urmează:

$$F_1(x) = \lim_{y \to +\infty} F(x,y),\ (\forall) x \in \mathbb{R}$$

$$F_2(y) = \lim_{x \to +\infty} F(x,y),\ (\forall) y \in \mathbb{R}$$

şi cum f este funcţie de densitate pentru variabila aleatoare bidimensională (X, Y) putem adăuga şi:

$$F(x,y) = \int_{-\infty}^{x}\int_{-\infty}^{y} f(x,y)\,dx\,dy \; ; \quad (\forall)(x,y) \in \mathbb{R}^2$$

Deci putem scrie:

$$F_1(x) = \int_{-\infty}^{x}\int_{-\infty}^{+\infty} f(x,y)\,dx\,dy \; ; \quad (\forall)x \in \mathbb{R}$$

$$F_2(y) = \int_{-\infty}^{+\infty}\int_{-\infty}^{y} f(x,y)\,dx\,dy \; ; \quad (\forall)y \in \mathbb{R}$$

ceea ce arată că funcţiile f_1, f_2 definite în enunţul teoremei sînt funcţii de densitate pentru variabilele aleatoare marginale X, Y, respectic şi teorema este demonstrată.

Pe lîngă variabilele aleatoare marginale, care se pot asocia unei variabile aleatoare bidimensionale, este uneori util să se considere şi alte categorii de variabile aleatoare unidimensionale asociate unei variabile aleatoare bidimensionale date. În continuare vom preciza ce se înţelege prin variabila aleatoare asociată unei variabile aleatoare bidimensionale, considerînd pentru început numai variabile aleatoare bidimensionale de tip discret.

Fie (X, Y) o variabilă aleatoare bidimensională de tip discret, definită peste un cîmp de probabilitate (Ω, K, P), avînd tabloul de repartiţie:

$$\begin{pmatrix} (x_i, y_i) \\ p_{ij} \end{pmatrix} \; ; \quad i \in I, \; j \in J$$

Dacă y_j este o valoare dată a variabilei aleatoare marginală Y atunci mulţimea $\Omega^j = \{\omega \mid Y(\omega) = y_j\} \in K$ iar din teorema 7.5.5. urmează:

$$P(\Omega^j) = P(Y = y_j) = \sum_{i \in I} p_{ij}$$

Considerînd acum $K^j = \{A \cap \Omega^j ; A \in K\}$ şi funcţia $P^j : K^j \to \mathbb{R}$ definită prin:

$$P^j(A) = \frac{P(A)}{P(\Omega^j)} \; ; \quad (\forall) A \in K^j$$

se constată imediat că tripletul (Ω^j, K^j, P^j) este un cîmp de probabilitate, probabilitatea P^j fiind de fapt restricţia la K^j a probabilităţii condiţionate de evenimentul Ω^j corespunzătoare cîmpului de probabilitate (Ω, K, P). Deoarece X este variabilă aleatoare peste (Ω, K, P), $X: \Omega \to \mathbb{R}$, rezultă că restricţia funcţiei X la mulţimea Ω^j, este o variabilă aleatoare peste cîmpul de probabilitate (Ω^j, K^j, P^j). Restricţia variabilei aleatoare X la mulţimea $\Omega^j \subset \Omega$ o vom nota $X \mid Y = y_i$ şi cum aceasta este o variabilă aleatoare o vom numi variabila aleatoare X condiţionată de $Y = y_j$. Observăm că, procedînd în acest mod, pentru fiecare valoare a variabilei marginale Y, y_i, se poate asocia o variabilă aleatoare condiţionată de forma $X \mid Y = y_j$. Dacă în raţionamentul de mai înainte schimbăm rolul variabilelor X, Y atunci, parcurgînd acelaşi drum, constatăm că putem defini şi variabile aleatoare condiţionate de forma $Y \mid X = x_i$, unde x_i este o valoare oarecare a variabilei marginale X. Să mai observăm că dacă (X, Y) este o variabilă aleatoare bidimensională de tip discret atunci variabilele aleatoare condiţionate construite ca mai înainte vor fi tot de tip discret deoarece mulţimea valorilor acestora va fi, în general, o submulţime a mulţimii valorilor variabilelor aleatoare marginale.

Să vedem acum ce legătură există între repartiţia variabilei aleatoare bidimensionale (X, Y) şi repartiţiile variabilelor aleatoare condiţionate. Continuînd să ne limităm la cazul variabilelor de tip discret, observăm că:

$$P^j((X \mid Y = y_j) = x_i) = \frac{P(X = x_i, Y = y_i)}{P(Y = y_i)} = \frac{p_{ij}}{q_j}, \; (\forall) i \in I$$

unde $q_j = \sum_{i \in I} p_{ij}$. Prin urmare variabilele aleatoare condiţionate $X \mid Y = y_j, j \in J$, au tabloul de repartiţie de forma:

$$\begin{pmatrix} x_i \\ \dfrac{p_{ij}}{q_j} \end{pmatrix}, \quad i \in I$$

unde $q_j = \sum_{i \in I} p_{ij}$ iar $p_{ij} = P(X = x_i, Y = y_j)$, $i \in I$, $j \in J$. În mod asemănător se arată că tabloul de repartiţie pentru o variabilă aleatoare condiţionată de forma $Y | X = x_i$ va fi:

$$\begin{pmatrix} y_i \\ p_{ij} \\ \hline p_i \end{pmatrix}; \quad j \in J$$

unde $p_i = \sum_{j \in J} p_{ij}$, iar $p_{ij} = P(X = x_i, Y = y_j)$, $i \in I$, $j \in J$.

Fie acum (X, Y) o variabilă aleatorie bidimensională de tip continuu cu funcţia de densitate $f: \mathbf{R}^2 \to \mathbf{R}^2$, variabilele aleatoare marginale X, Y avînd funcţiile de densitate f_1 şi f_2, respectiv. Dacă $y \in \mathbf{R}$ este astfel încît $f_2(y) \neq 0$, generalizînd consideraţiile de mai înainte, relativ la cazul discret, se constată că este natural ca prin variabilă aleatoare condiţionată $X | Y = y$ să se înţeleagă variabila aleatoare unidimensională a cărei funcţie de densitate este $\varphi_y : \mathbf{R} \to \mathbf{R}$, unde:

$$\varphi_y(x) = \frac{f(x, y)}{f_2(y)}, \quad (\forall) x \in \mathbf{R}$$

În mod asemănător, dacă $x \in \mathbf{R}$ este astfel încît $f_1(x) \neq 0$, prin variabila aleatoare condiţionată se înţelege variabila aleatoare unidimensională a cărei funcţie de densitate este $\Psi_x : \mathbf{R} \to \mathbf{R}$, definită prin:

$$\Psi_x = \frac{f(x, y)}{f_1(x)}; \quad (\forall) y \in R$$

Rezumăm consideraţiile de pînă acum în următoarele două definiţii:

Definiţia 7.5.14.

Fie (X, Y) o variabilă aleatoare bidimensională de tip discret, cu tabloul de repartiţie:

$$\begin{pmatrix} (x_i, y_i) \\ p_{ij} \end{pmatrix}; \quad i \in I, j \in J$$

Dacă y_i este o valoare fixată a variabilei aleatoare marginale Y, iar x_i o

valoare fixată a variabilei aleatoare marginale X, atunci prin variabilele aleatoare condiţionate $X|Y=y_i$ şi $Y|X=x_i$ se înţeleg variabilele aleatoare unidimensionale cu tablourile de repartiţie următoare:

$$X|Y=y_i:\begin{pmatrix} x_i \\ \hline p_{ij} \\ \hline q_j \end{pmatrix}, i\in I\ ;\ Y|X=x_i:\begin{pmatrix} y_j \\ \hline p_{ij} \\ \hline p_i \end{pmatrix}, j\in J$$

unde $q_j=\sum_{i\in I} p_{ij}$, $p_i=\sum_{j\in J} p_{ij}$

Definiţia 7.5.15.

Fie (X,Y) o variabilă aleatoare bidimensională cu funcţia de densitate $f:\mathbb{R}^2\to\mathbb{R}^2$. Dacă funcţiile $f_1, f_2:\mathbb{R}\to\mathbb{R}$ sînt funcţiile de densitate ale variabilelor aleatoare marginale X, Y, respectiv, iar $x_0, y_0\in\mathbb{R}$ sînt astfel încît $f_1(x_0)\neq 0$, $f_2(y_0)\neq 0$, prin variabilele aleatoare condiţionate $X|Y=y_0$, $Y|X=x_0$ se înţeleg variabilele aleatoare unidimensionale ale căror funcţii de densitate φ_{y_0}, Ψ_{x_0} sînt definite prin:

$$\varphi_{y_0}=\frac{f(x,y_0)}{f_2(y_0)},\ (\forall)x\in\mathbb{R}$$

$$\Psi_{x_0}=\frac{f(x_0,y)}{f_1(x_0)},\ (\forall)y\in\mathbb{R}$$

Exerciţii:

1. Se dă cîmpul de probabiliate (Ω,K,P) în care $\Omega=\{\omega_1,\omega_2,\omega_3,\omega_4\}$ evenimentele elementare din K fiind $A_1=\{\omega_1,\omega_2\}$, $A_2=\{\omega_3\}$, $A_3=\{\omega_4\}$, iar probabilitatea P definită prin $P(A_1)=\frac{1}{2}$, $P(A_2)=\frac{1}{4}$, $P(A_3)=\frac{1}{4}$. Să se arate că funcţia $X:\Omega\to\mathbb{R}$, definită prin $X(\omega_1)=X(\omega_2)=0$, $X(\omega_3)=-1$, $X(\omega_4)=\frac{1}{2}$ este o variabilă aleatoare peste cîmpul de probabilitate (Ω,K,P). Să se construiască funcţia de repartiţie a acestei variabile aleatoare.

2. Să se arate că, dacă X, Y sînt două variabile aleatoare independente, definite peste un cîmp de probabilitate (Ω, K, P), atunci

a) $P(X<x, Y\geq y) = P(X<x)\,P(Y\geq y)$, $(\forall)\,x, y\in\mathbb{R}$

b) $P(X\geq x, Y<y) = P(X\geq x)\,P(Y<y)$, $(\forall)\,x, y\in\mathbb{R}$

c) $P(X\geq x, Y\geq y) = P(X\geq x)\,P(Y\geq y)$, $(\forall)\,x, y\in\mathbb{R}$

3. Să se arate că, în condiţiile exerciţiului precedent, dacă are loc una din relaţiile a), b), c) atunci variabilele aleatoare X, Y sînt independente.

4. Fie X o variabilă aleatoare bidimensională definită peste un cîmp borelian de probabilitate (Ω, K, P), astfel că $X(\omega) = (X_1(\omega), X_2(\omega))$ pentru oricare ω, $\omega\in\Omega$. Să se arate că funcţiile $X_1 : \Omega\to\mathbb{R}$, $X_2 : \Omega\to\mathbb{R}$ sînt variabile aleatoare unidimensionale peste cîmpul de probabilitate (Ω, K, P).

5. Se consideră tabelul:

$$\begin{pmatrix} 0 & 1 & 2 \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} \end{pmatrix}$$

Este posibil ca acest tabel să constituie tabloul de repartiţie al unei variabile aleatoare discrete?

6. Se consideră tabelul:

$$\begin{pmatrix} -1 & 0 & 1 & 3 \\ \frac{c}{2} & \frac{c}{3} & \frac{c}{4} & \frac{c}{5} \end{pmatrix}$$

în care c este un număr oarecare. Pentru care valoare a lui c acest tabel poate constitui tabloul de repartiţie al unei variabile aleatoare discrete?

7. Să se reprezinte grafic funcţia de repartiţie a variabilei aleatoare discrete avînd tabloul de repartiţie următor:

$$\begin{pmatrix} -2 & 1 & 2 & 5 \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{4} & \frac{1}{6} \end{pmatrix}$$

8. Presupunem că X este o variabilă aleatoare de tip continuu, cu densitatea de probabilitate f, să se arate că are loc:

$$P(a<X<b)=\int_a^b f(t)dt$$

9. Fie $X_1, X_2, \ldots, X_n$ n variabile de tip discret, definite peste un cîmp de probabilitate (Ω, K, P). Să se arate că variabilele aleatoare X_i, $i=1,\ldots,n$, sînt independente dacă şi numai dacă pentru oricare numere reale x_i, $i=1,\ldots,n$, astfel încît x_i să fie o valoare a variabilelor aleatoare X_i, $1 \le i \le n$, are loc relaţia:

$$P(X_1=x_1, X_2=x_2, \ldots, X_n=x_n)=P(X_1=x_1)\,P(X_2=x_2)\ldots P(X_n=x_n)$$

10. Fie (Ω, K, P) un cîmp de probabilitate în care $\Omega=\{\omega_1, \omega_2\}$, $K=P(\Omega)$ şi fie $X_1:\Omega\to\mathbb{R}$, $X_2:\Omega\to\mathbb{R}$ două funcţii definite prin $X_1(\omega_1)=0$, $X_1(\omega_2)=1$, $X_2(\omega_1)=1$, $X_2(\omega_2)=0$. Ce condiţii ar trebui să îndeplinească probabilitatea P pentru ca funcţiile X_1, X_2, considerate ca variabile aleatoare peste cîmpul de probabilitate (Ω, K, P), să fie variabile aleatoare independente.

11. Se dă funcţia $f:\mathbb{R}\to\mathbb{R}$, definite prin:

$$f(x)=\begin{cases}\dfrac{1}{n}e^{-\frac{x}{n}}; & x>0\\ 0 \quad\quad ; & x\le 0\end{cases}$$

în care n este un număr natural oarecare. Să se arate funcţia f este o densitate de probabilitate şi să se determine funcţia de repartiţie a variabilei aleatoare a căror funcţie de densitate este f. Ce se poate spune despre repartiţia valorilor variabilei aleatoare cu funcţia f atunci cînd n este foarte mare?

12. Fie X o variabilă aleatoare de tip continuu, a cărei funcţie de densitate este f, $f:\mathbb{R}\to\mathbb{R}$. Să se arate că variabila aleatoare X^{2k}, $k\ge 1$, are funcţia de densitate $\varphi_1:\mathbb{R}\to\mathbb{R}$, definită prin:

$$\varphi_1(x)=\begin{cases}0 & ; x\le 0\\ \dfrac{1}{2k}x^{\frac{1}{2k}-1}\left[f(x^{\frac{1}{2k}})+f(-x^{\frac{1}{2k}})\right] & ; x>0\end{cases}$$

iar variabila aleatoare X^{2k+1}, $k\ge 1$, are funcţia de densitate $\varphi_2:\mathbb{R}\to\mathbb{R}$, definită prin:

$$\varphi_2(x) = \frac{1}{2k+1} x^{-\frac{2k}{2k+1}} f(x^{\frac{1}{2k}+1}), \ (\forall) x \in \mathbb{R}$$

13. Fie (Ω, K, P) un cîmp de probabilitate şi X, Y două variabile aleatoare de tip discret, independente, definite peste acest cîmp de probabilitate, avînd tablourile de repartiţie:

$$X: \begin{pmatrix} -1 & 0 & 1 \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \end{pmatrix} \ ; \ Y: \begin{pmatrix} 0 & 1 \\ \frac{1}{4} & \frac{3}{4} \end{pmatrix}$$

Să se construiască tablourile de repartiţie ale variabilelor aleatoare $X+Y$, $X \cdot Y$, X^2, $X-Y$, $3X$, $Y+2$, $\frac{X}{Y+2}$.

14. Se consideră un cîmp de probabilitate (Ω, K, P), în care $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4\}$, $K = \{\{\omega_1, \omega_2\}, \{\omega_3, \omega_4\}, \Phi, \Omega\}$, $P(\{\omega_1, \omega_2\}) = \frac{1}{3}$, $P(\{\omega_3, \omega_4\}) = \frac{2}{3}$, $P(\Phi) = 0$, $P(\Omega) = 1$.

Să se arate că funcţiile $X: \Omega \to \mathbb{R}$, $Y: \Omega \to \mathbb{R}$, definite prin $X(\omega_1) = X(\omega_2) = 1, X(\omega_3) = X(\omega_4) = 0, Y(\omega_1) = Y(\omega_2) = 1, Y(\omega_3) = Y(\omega_4) = -1$, sînt variabile aleatoare peste cîmpul de probabilitate (Ω, K, P) şi să se construiască tablourile de repartiţie ale variabilelor aleatoare, $X, Y, X+Y, \frac{X}{Y}, X \cdot Y$.

15. Fie X o variabilă aleatoare de tip discret cu tabloul de repartiţie:

$$\begin{pmatrix} x_i \\ p_i \end{pmatrix}, \ i = 1, 2, .., n$$

iar Y o variabilă aleatoare de tip continuu cu funcţia de densitate f, $f: \mathbb{R} \to \mathbb{R}$, (ambele variabile aleatoare se presupun definite peste un acelaşi cîmp de probabilitate). Să se arată că, dacă X, Y sînt variabile aleatoare independente, atunci variabila aleatoare $X+Y$ este o variabilă aleatoare de tip continuu avînd funcţia de densitate φ, $\varphi: \mathbb{R} \to \mathbb{R}$, definită prin:

$$\varphi(x) = \sum_{i=1}^{n} p_i f(x - x_i), \ (\forall) x \in \mathbb{R}$$

16. Fie (X, Y) o variabilă aleatoare bidimensională de tip discret, cu tabloul de repartiţie:

$$\begin{pmatrix} (-1,0) & (1,-1) & (-1,2) & (1,2) & (1,1) & (-1,1) \\ \frac{1}{3} & \frac{1}{6} & \frac{1}{6} & \frac{1}{12} & \frac{1}{6} & \frac{1}{12} \end{pmatrix}$$

Să se determine tablourile de repartiţie pentru variabilele aleatoare marginale şi pentru variabilele aleatoare condiţionate, asociate variabilelor bidimensionale (X, Y).

Răspuns:

$$X: \begin{pmatrix} -1 & 1 \\ \frac{7}{12} & \frac{5}{12} \end{pmatrix}, \quad Y: \begin{pmatrix} -1 & 0 & 1 & 2 \\ \frac{1}{6} & \frac{1}{3} & \frac{1}{4} & \frac{1}{4} \end{pmatrix}$$

$$X \mid Y=2: \begin{pmatrix} -1 & 1 \\ \frac{2}{3} & \frac{1}{3} \end{pmatrix}, \quad X \mid Y=1: \begin{pmatrix} -1 & 1 \\ \frac{1}{3} & \frac{2}{3} \end{pmatrix}$$

$$X \mid Y=0: \begin{pmatrix} -1 \\ 1 \end{pmatrix}, \quad X \mid Y=-1: \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

$$Y \mid X=-1: \begin{pmatrix} 0 & 1 & 2 \\ \frac{4}{7} & \frac{1}{7} & \frac{2}{7} \end{pmatrix}, \quad Y \mid X=1: \begin{pmatrix} -1 & 1 & 2 \\ \frac{2}{5} & \frac{2}{5} & \frac{1}{5} \end{pmatrix}$$

17. Fie (X, Y) o variabilă aleatoare bidimensională de tip discret definită peste un cîmp de probabilitate (Ω, K, P) şi $y_1 \in Y(\Omega)$ aşa fel încît $P(Y=y_i) = P(\Omega_j) \neq 0$, unde $\Omega_j = \{\omega \in \Omega ; Y(\omega) = y_j\}$. Dacă pentru $A \in K$ definim $P^j(A) = P(A \mid \Omega_j)$ atunci P^j este o probabilitate peste (Ω, K) şi putem considera cîmpul de probabilitate (Ω, K, P^j), unde $X^j(\omega) = X(\omega)$, $(\forall)\, \omega \in \Omega$. Să se arate că variabilele aleatoare X^j şi $X \mid Y = y_i$ au aceleaşi funcţii de repartiţie.

§ 6. CARACTERISTICI NUMERICE ALE UNEI VARIABILE ALEATOARE

După cum am mai menţionat, în problemele practice se întîlnesc deseori variabile care iau valori în funcţie de producerea unuia sau mai multor evenimente dintr-un cîmp dat de evenimente. Considerînd o astfel de variabilă, este util să fie puse în evidenţă anumite proprietăţi ale sale prin determinarea unor indicatori numerici sintetici, asociaţi variabilei care ne interesează. Astfel de indicatori urmăresc să pună în evidenţă pe de o parte tendinţa centrală de grupare a valorilor variabilei respective, iar pe de altă parte gradul de împrăştiere a valorilor variabilei. În acest paragraf vom prezenta cele mai des utilizate caracteristici numerice ale unei variabile aleatoare, împreună cu proprietăţile cel mai importante ale acestora, iar în paragraful următor vor fi determinate aceste caracteristici numerice unele repartiţii frecvent întîlnite în practică.

Media unei variabile aleatoare

Deoarece pe parcursul celor ce urmează vom întîlni numai variabile aleatoare de tip discret de tip continuu, vom defini media, ca de altfel şi celelalte caracteristici numerice ale unei variabile aleatoare, numai pentru variabile aleatoare de tip discret sau de tip continuu.

Definiţia 7.6.1.

Fie X o variabilă aleatoare discretă, definită peste un cîmp de probabilitate (Ω, K, P), care poate lua valorile $x_i, i = 1, 2, \ldots,$, cu probabilităţile corespunzătoare $p_i, i = 1, 2, \ldots$. Prin media acestei variabile aleatoare (sau valoare medie a variabilei aleatoare X) înţelegem numărul $M(X)$, definit prin:

$$M(X) = \sum_{i=1}^{n} x_i p_i \qquad 7.6.1$$

dacă variabila aleatoare X are un număr finit de valori, x_i, $i = 1, 2, \ldots, n$ şi prin

$$M(X) = \sum_{i=1}^{\infty} x_i p_i \qquad 7.6.1'$$

atunci cînd variabila aleatoare X are o mulţime numărabilă de valori, $x_i, i = 1, 2, \ldots$, cu condiţia ca seria $\sum_{i=1}^{\infty} x_i p_i$ să fie convergentă.

În cazul cînd o variabilă discretă X nu este simplă, iar seria $\sum_{i=1}^{\infty} x_i p_i$ nu este convergentă, atunci vom spune că variabila aleatoare X nu are medie. Pentru exemplificare să considerăm variabila aleatore discretă X cu tabloul de repartiţie:

$$\begin{pmatrix} 1 & 2 & 2^2 & \ldots & 2^n & \ldots \\ \frac{1}{2} & \frac{1}{2^n} & \frac{1}{2^3} & \ldots & \frac{1}{2^{n+1}} & \ldots \end{pmatrix}$$

Se poate verifica imediat că acest tablou este tablou de reapartiţie al unei variabile aleatoare, dar această variabilă aleatoare nu are medie deoarece:

$$\sum_{i=1}^{\infty} x_i p_i = \sum_{i=1}^{\infty} 2^i \cdot \frac{1}{2^{i+1}} = \sum_{i=1}^{\infty} \frac{1}{2}$$

ceea ce arată că seria $\sum_{i=1}^{\infty} x_i p_i$ nu este convergentă.

Definiţia 7.6.2.

Fie X o variabilă aleatoare de tip continuu, a cărei funcţie de densitate este f. Dacă integrala:

$$\int_{-\infty}^{+\infty} x f(x)\, dx$$

este convergentă atunci prin media variabilei aleatoare X, $M(X)$, se înţelege valoarea acestei integrale, adică

$$M(X) = \int_{-\infty}^{+\infty} x f(x) dx \qquad 7.6.2$$

Valoarea medie a unei variabile aleatoare poate fi interpretată ca fiind un număr „central" în mulţimea valorilor variabilei aleatoare, de fapt care va

fi relevat atît prin exemplele pe care le vom considera în paragraful următor, cît şi prin consideraţiile din partea de statistică matematică. De altfel, un prim argument al faptului că media unei variabile aleatoare reprezintă, într-un anume sens, un număr care se află „în centrul" valorilor variabilei aleatoare îl vom prezenta după ce vom pune în evidenţă cîteva proprietăţi importante ale mediei unei variabile aleatoare.

Să convenim de la început ca prin *constantă* să înţelegem o variabilă aleatoare (simplă) care poate lua o singură valoare, iar dacă această valoare este C atunci variabila aleatoare repectivă să o notăm tot prin C.

Se poate dovedi că media unei variabile aleatoare are următoarele proprietăţi:

1. Dacă C este o constantă, atunci $M(C) = C$.

2. Dacă $c \in \mathbf{R}$, iar X este o variabilă aleatoare, atunci $M(cX) = cM(X)$.

3. Dacă X, Y sînt două variabile aleatoare definite peste acelaşi cîmp de probabilitate, atunci:

$$M(X + Y) = M(X) + M(Y)$$

4. Dacă X, Y sînt două variabile aleatoare independente, atunci:

$$M(X \cdot Y) = M(X)\,M(Y)$$

Vom justifica aici doar primele două dintre aceste proprietăţi, justificarea celorlalte două proprietăţi putînd fi găsită în [9].

Demonstraţie:

1. Dacă interpretăm o constantă c, ca o variabilă aleatoare simplă ce poate lua o singură valoare c,

$$\begin{pmatrix} c \\ 1 \end{pmatrix}$$

şi conform definiţiei 7.6.1, media acesteia va fi $c \cdot 1 = c$, ceea ce arată că $M(c) = c$.

2. a) Considerînd $c \in \mathbf{R}$ şi X o variabilă aleatoare discretă cu tabloul de repartiţie:

$$\begin{pmatrix} x_i \\ p_i \end{pmatrix}, \ i = 1, 2, \ldots$$

atunci variabila aleatoare cX va avea tabloul de repartiţie:

$$\begin{pmatrix} c.x_i \\ p_i \end{pmatrix}, \ i = 1, 2, \ldots$$

Folosind acum definiţia 7.6.1, rezultă că, atît în cazul cînd variabila X este simplă cît şi în cazul variabila aleatoare X are o mulţime numărabilă de valori şi are medie, putem scrie

$$M(X) = \sum_i x_i p_i$$

$$M(cX) = \sum_i (cx_i) p_i = c \sum_i x_i p_i = cM(X)$$

Prin urmare $M(cX) = cM(X)$, ceea ce arată că proprietatea 2 este adevărată în cazul cînd X este o variabilă aleatoare discretă.

b) Fie acum $c \in \mathbb{R}$ şi X o variabilă aleatoare de tip continuu, avînd funcţia de densitate f. Este clar că atunci cînd $c = 0$ urmează $M(cX) = 0 = cM(X)$ şi deci putem presupune $c \neq 0$. După cum am constatat în paragraful precedent, funcţia de densitate a variabilei aleatoare cX, $\varphi : \mathbb{R} \to \mathbb{R}$, este dată prin:

$$\varphi(x) = \frac{1}{|c|} f\left(\frac{x}{c}\right), \ (\forall) x \in \mathbb{R}$$

şi atunci putem scrie:

$$M(cX) = \int_{-\infty}^{+\infty} x . \frac{1}{|c|} f\left(\frac{x}{c}\right) dx$$

Dacă în integrala din membrul al doilea al relaţiei precedente se face substituţia $\frac{x}{c} = t$ urmează $dx = cdt$ şi cum $|c| = c$ dacă $c > 0$, iar $|c| = -c$

dacă $c<0$ rezultă:

$$M(cX)=\begin{cases} c\int_{-\infty}^{+\infty} tf(t)dt\,, & c>0 \\ -c\int_{+\infty}^{-\infty} tf(t)dt\,, & c<c \end{cases}$$

Dar $-\int_{+\infty}^{-\infty} tf(t)dt=\int_{-\infty}^{+\infty} tf(t)dt$ şi atunci urmează:

$$M(cX)=c\int_{-\infty}^{+\infty} tf(t)dt$$

atît în cazul $c>0$, cît şi în cazul $c<0$. Totodată, conform definiţiei 7.6.2,

$$M(X)=\int_{-\infty}^{+\infty} tf(t)dt$$

şi deci $M(cX)=cM(X)$ şi cazul cînd X este o variabilă aleatoare de tip continuu, ceea ce completează demonstraţia proprietăţii 2.

Să vedem acum cum s-ar putea justifica denumirea de „valoare centrală", între valorile unei variabile aleatoare, pentru media variabilei aleatoare considerate.

Să presupunem că X este o variabilă aleatoare iar $M(X)$ este media sa, care, conform definiţiei, este un număr real. Peste cîmpul de probabilitate (Ω, K, P), peste care este definită variabila aleatoare X, putem considera o nouă variabilă aleatoare $X-M(X)=Y$ definită prin:

$$Y(\omega)=X(\omega)-M(X)\,,\quad Y=X-M(X)\quad (\forall)\,\omega\in\Omega$$

de unde se vede că valorile variabilei aleatoare $Y=X-M(X)$ sînt diferenţele dintre valorile variabilei aleatoare X şi media sa $M(X)$, aceste diferenţe fiind pozitive sau negative după cum valorile corespunzătoare ale variabilei aleatoare X sînt mai mari sau mai mici decît media $M(X)$. Din acest motiv variabila aleatoare $X-M(X)$ se mai numeşte şi *abatere a variabilei aleatoare X de la media sa*. Considerînd acum media variabilei $Y=X-M(X)$, folosind proprietăţile mediei enunţate mai sus, obţinem:

$$M(X - M(X)) = M(X) - M(M(X)) = M(X) - M(X) = 0$$

ceea ce arată că media variabilei aleatoare abatere a lui de la media sa este egală cu zero. Acest fapt justifică acum denumirea de valoare centrală pentru media unei variabile aleatoare, deoarece $M(X - M(X)) = 0$ ne sugerează că abaterile variabilei aleatoare X la stînga şi la dreapta mediei sale se compensează reciproc.

Dispersia unei variabile aleatoare

Am făcut mai înainte observaţia că media unei variabile aleatoare este un număr, într-un anumit sens „central" între valorile pe care le poate lua variabila aleatoare respectivă. Este de o deosebită importanţă practică cunoaşterea gradului de împrăştiere a valorilor unei variabile aleatoare asociate unei anumite caracteristici a unei colectivităţi. O măsură frecvent utilizată a gradului de împrăştiere a valorilor unei variabile aleatoare este *dispersia* variabilei aleatoare respective, aceasta evaluînd gradul de împrăştiere a valorilor variabilei aleatoare faţă de media sa. (Denumirea de *dispersie*, care înseamnă împrăştiere, provine de la cuvîntul latinesc *dispersio*).

Definiţia 7.6.3.

Fie X o variabilă aleatoare definită peste un cîmp de probabilitate (Ω, K, P). Prin dispersia variabilei aleatoare X înţelegem numărul $D(X)$, definit prin:

$$D(X) = M[(X - M(X))^2] \qquad 7.6.3$$

adică dispersia unei variabile aleatoare se defineşte ca fiind media pătratului variabilei aleatoare abatere a variabilei aleatoare date de la media sa. Ţinînd cont de această definiţie, se poate spune că dispersia unei variabile aleatoare constituie o măsură a gradului de împrăştiere a valorilor variabilei aleatoare faţă de media sa. Pe baza relaţiei de definiţie a dispersiei unei variabile aleatoare X, putem constata imediat că are loc relaţia:

$$D(X) = M(X^2) - [M(X)]^2 \qquad 7.6.4$$

Într-adevăr, pe baza proprietăţilor medie unei variabile aleatoare şi a definiţiei

dispersiei, putem scrie succesiv:

$$D(X) = M[(X - M(X))^2] = M[X^2 - 2XM(X) + [M(X)]^2] =$$
$$= M(X^2) - 2M(X)M(X) + [M(X)]^2 = M(X^2) - [M(X)]^2$$

ceea ce justifică realţia 7.6.4.

Vom dovedi în continuare că dispersia unei variabile aleatoare are următoarele proprietăţi:

1. Dacă C este o constantă atunci $D(C) = 0$.
2. Dacă $c \in \mathbf{R}$ şi X este o variabilă aleatoare atunci:

$$D(cX) = c^2 D(X)$$

3. Dacă $c \in \mathbf{R}$ şi X este o variabilă aleatoare atunci

$$D(c + X) = D(X)$$

4. Dacă X şi Y sînt două variabuile aleatoare independente, definite peste acelaşi cîmp de probabilitate, atunci

$$D(X + Y) = D(X) + D(Y)$$

Demonstraţie:

1. Deoarece $M(C) = C$, din relaţia de definiţie a dispersiei unei variabile aleatoare, urmează:

$$D(C) = M((C - C)^2) = M(0) = 0$$

2. Fie $c \in \mathbf{R}$ şi X o variabilă aleatoare. Din definiţia dispersiei şi proprietatea 2 a mediei unei variabile aleatoare putem scrie succesiv:

$$D(cX) = M[(cX - M(cX))^2] = M[c^2(X - M(X))^2] =$$
$$= c^2 M[(X - M(X))^2] = c^2 D(X)$$

ceea ce demonstrează proprietatea 2 a dispersiei unei variabile aleatoare.

3. Fie $c \in \mathbf{R}$ şi X o variabilă aleatoare. Atunci, pe baza proprietăţilor mediei unei variabile aleatoare, putem scrie succesiv:

$$D(c + X) = M[(c + X - M(c + X))^2] = M[(c + X - C - M(X))^2] =$$
$$= M[(X - M(X))^2] = D(X)$$

proprietatea 3 a dispersiei este demonstrată.

4. Fie X, Y două variabile aleatoare independente definite peste un acelaşi cîmp de probabilitate (Ω, K, P). Pe baza proprietăţilor mediei unei variabile şi a definiţiei dispersiei rezultă că putem scrie succesiv:

$$D(X+Y) = M\left[(X+Y-M(X+Y))^2\right] = M\left[(X-M(X))^2\right] + \\ + M\left[(Y-M(Y))^2\right] + 2M\left[(X-M(X))(Y-M(Y))\right]$$

deoarece din independenţa variabilelor aleatoare X, Y urmează că şi variabilele aleatoare $X-M(X)$, $Y-M(Y)$ sînt de asemeni independente.

Dar $M(X-M(X)) = 0$ şi deci, din relaţiile de mai sus urmează:

$$D(X+Y) = D(X) + D(Y)$$

ceea ce arată că proprietatea 4 a dispersiei este adevărată.

Atunci cînd am definit media unei variabile aleatoare am precizat implicit şi modul în care aceasta se determină, în cazul variabilelor aleatoare de tip discret şi de tip continuu.

Dispersia unei variabile aleatoare a fost definită ca media unei variabile aleatoare asociate variabilei aleatoare date, (media pătratului variabilei aleatoare abatere de la medie), modul de calcul al dispersiei unei variabile aleatoare de tip discret sau de tip continuu putînd fi dedus direct din modul de calcul al mediei unei variabile aleatoare de tip discret sau de tip continuu.

Vom arăta că atunci cînd X este o variabilă aleatoare de tip discret, cu tabloul de repartiţie de forma:

$$\begin{pmatrix} x_i \\ p_i \end{pmatrix}, \; i = 1, 2, \ldots$$

atunci:

$$D(X) = \sum_i (x_i - m)^2 p_i \qquad 7.6.5$$

unde prin m am notat media variabilei aleatoare X, adică $m = M(X)$.

Totodată, în cazul în care X este o variabilă aleatoare de tip continuu, cu funcţia de densitate f, vom arăta că are loc relaţia:

$$D(X) = \int_{-\infty}^{+\infty} (x-m)^2 f(x)\,dx \qquad 7.6.6$$

unde m notează, ca şi în cazul precedent, media variabilei aleatoare X, adică

$m = M(X)$. Într-adevăr, dacă X este o variabilă aleatoare de tip discret, cu tabloul de repartiţie de forma:

$$\begin{pmatrix} x_i \\ p_i \end{pmatrix}, \; i = 1, 2, \ldots$$

atunci variabila aleatoare $(X - M(X))^2 = (X - m)^2$ este de asemeni o variabilă aleatoare de tip discret, al cărui tablou de repartiţie va fi:

$$\begin{pmatrix} (x_i - m)^2 \\ p_i \end{pmatrix}, \; i = 1, 2, \ldots$$

Pe baza definiţiei dispersiei şi a mediei unei variabile aleatoare de tip discret rezultă:

$$D(X) = \sum_i (x_i - m)^2 p_i$$

ceea ce justifică relaţia (7.6.5).

Fie acum o variabilă aleatoare de tip continuu, X, a cărei funcţie de densitate este $f, f: \mathbb{R} \to \mathbb{R}$. Din teorema 7.5.2 rezultă atunci că funcţia de densitate f_1 a variabilei aleatoare $X - M(X) = X - m$, poate fi definită prin:

$$f_1(x) = f(x + m), \; x \in \mathbb{R}$$

iar funcţia de densitate φ a variabilei aleatoare $(X - M(X))^2$, poate fi definită prin:

$$\varphi(x) = \begin{cases} \dfrac{1}{2\sqrt{x}} f_1(\sqrt{x}) + f_1(-\sqrt{x}), & x > 0 \\ 0, & x \le 0 \end{cases}$$

Deci putem scrie:

$$D(X) = \int_{-\infty}^{+\infty} x\,\varphi(x)\,dx = \int_0^{\infty} x\,\varphi(x)\,dx$$

deoarece $\varphi(x) = 0$ pentru $x \le 0$.

Pe de altă parte, dacă $x > 0$ atunci:

$$\varphi(x) = \frac{1}{2\sqrt{x}} f_1(\sqrt{x}) + f_1(-\sqrt{x}) = \frac{1}{2\sqrt{x}} f(\sqrt{x} + m) + f(-\sqrt{x} + m)$$

BCU IASI / CENTRAL UNIVERSITY LIBRARY

care împreună cu relaţia de mai sus arată că

$$D(X) = \int_0^{\infty} x \frac{1}{2\sqrt{x}} f(\sqrt{x}+m)dx + \int_0^{\infty} x \frac{1}{2\sqrt{x}} f(-\sqrt{x}+m)dx$$

Dacă în prima dintre aceste integrale se face substituţia $\sqrt{x}+m=t$, iar în cea de a doua integrală substituţia $-\sqrt{x}+m=t$ se obţine:

$$D(X) = \int_0^{\infty} (t-m)^2 f(t)dt + \int_{-\infty}^{0} (t-m)^2 f(t)dt$$

adică:

$$D(X) = \int_{-\infty}^{+\infty} (t-m)^2 f(t)dt$$

ceea ce arată că relaţia (7.6.6) este adevărată.

Exemplul 7.6.1.

a) Fie X o variabilă aleatoare de tip discret, al cărei tablou de repartiţie este:

$$\begin{pmatrix} -1 & 0 & 1 & 3 \\ \frac{1}{4} & \frac{2}{5} & \frac{1}{10} & \frac{1}{4} \end{pmatrix}$$

În această situaţie media variabilelor aleatoare $X, M(X)$, va fi:

$$M(X) = (-1)\frac{1}{4} + 0 \cdot \frac{2}{5} + 1 \cdot \frac{1}{10} + 3 \cdot \frac{1}{4}$$

adică $M(X) = 0{,}6$, iar dispersia acestei variabile aleatoare, $D(X)$, va fi:

$$D(X) = (-1-\frac{3}{5})^2 \frac{1}{4} + (0-\frac{3}{5})^2 \frac{2}{5} + (1-\frac{3}{5})^2 \frac{1}{10} + (3-\frac{3}{5})^2 \frac{1}{4}$$

adică $D(X) = 2{,}24$.

b) Să considerăm acum o variabilă aleatoare de tip continuu, X, a cărei funcţie de densitate $f, f: \mathbb{R} \to \mathbb{R}$, este dată prin:

$$f(x) = \begin{cases} x + \frac{1}{2} & ; x \in [0,1] \\ 0 & ; x \notin [0,1] \end{cases}$$

BCU "LUCIAN BLAGA" CENTRAL UNIVERSITY LIBRARY

În această situaţie media variabilei aleatoare X este:

$$M(X) = \int_0^1 x(x + \frac{1}{2})dx = \frac{7}{12}$$

iar dispersia variabilei aleatoare X este:

$$D(X) = \int_0^1 (x - \frac{7}{12})^2 (x + \frac{1}{2})dx = \frac{11}{144}$$

Dispersia unei variabile aleatoare X, ca măsură a împrăştierii valorilor variabilei aleatoare date faţă de media sa, a fost definită ca media pătratului variabilei aleatoare abatere, $X - M(X)$. Prin ridicarea la pătrat a variabilei aleatoare $X - M(X)$ se schimbă unitatea de măsură pentru abateri şi atunci, pentru a se reveni la unitatea de măsură iniţială, drept indicator al împrăştierii valorilor variabilei aleatoare X se poate considera numărul $\sqrt{D(X)}$, numit şi *abatere medie pătratică* a variabilei aleatoare X. De obicei abaterea medie pătratică a unui variabile abaterea X se notează cu σ_x.

Variabile aleatoare normate

Fie X o variabilă aleatoare definită peste un cîmp de probabilitate $(\Omega, K, \mathcal{P})$ şi fie $M(X)$ şi $D(X)$ media şi dispersia acestei variabile aleatoare. Dacă $D(X) \neq 0$ atunci putem asocia variabilei aleatoare X o altă variabilă aleatoare, Y, definită peste acelaşi cîmp de probabilitate, unde:

$$Y = \frac{X - M(X)}{\sqrt{D(X)}}$$

Variabila aleatoare Y astfel construită se numeşte *variabila aleatoare normată* asociată variabilei aleatoare X. Vom arăta că pentru orice variabilă aleatoare X, care are dispersie $D(X) \neq 0$ variabila aleatoare normată asociată ei are media egală cu zero şi dispersia egală cu 1. Într-adevăr, ţinînd cont de proprietăţile mediei şi dispersiei unei variabile aleatoare, putem scrie:

$$M\left(\frac{X-M(X)}{\sqrt{D(X)}}\right)=\frac{1}{\sqrt{D(X)}}M(X-M(X))=0$$

$$D\left(\frac{X-M(X)}{\sqrt{D(X)}}\right)=\frac{1}{D(X)}D(X-M(X))=\frac{1}{D(X)}D(X)=1$$

ceea ce arată că pentru variabila aleatoare normată, asociată unei variabile aleatoare date, media este zero iar dispersia este 1.

Momente iniţiale şi momente centrate ale unei variabile aleatoare

Pe lîngă caracteristicile numerice ale unei variabile aleatoare prezentate pînă acum, (media şi dispersia), unei variabile aleatoare i se pot asocia şi alti indicatori numerici, caracteristici pentru repartiţia variabilei aleatoare respective. Între aceştia, momentele iniţiale şi momentele centrate de diferite ordine au un rol deosebit în analiza repartiţiei unei variabile aleatoare date.

Definiţia 7.6.4.

Fie X o variabilă aleatoare definită peste un cîmp de probabilitate $(\Omega, K, \mathcal{P})$. Numim moment iniţial de ordin k, $k \in \mathbb{N}$, pentru variabila aleatoare X, media variabilei aleatoare X^k.

Vom nota momentul iniţial de ordin k, al unei variabile aleatoare X prin $\alpha_k(X)$ şi atunci, conform definiţiei:

$$\alpha_k(X)=M(X^k)$$

Să observăm că media unei variabile aleatoare X, poate fi interpretată ca moment iniţial de ordin 1 al variabilei aleatoare, adică $a_1(X)=M(X)$.

Vom arăta că pentru o variabilă aleatoare de tip discret, X, cu tabloul de repartiţie:

$$\begin{pmatrix} x_i \\ p_i \end{pmatrix};\ i=1,2,\ldots$$

are loc:

$$M(X^k) = \sum_i x_i^k p_i \ , \quad k \in \mathbb{N} \qquad 7.6.7$$

iar pentru o variabilă aleatoare de tip continuu, cu funcţia de densitate $f, f: \mathbb{R} \to \mathbb{R}$, are loc:

$$M(X^k) = \int_{-\infty}^{+\infty} x^k f(x)dx \ , \ k \in \mathbb{N} \qquad 7.6.7'$$

a) Să presupunem că X este o variabilă aleatoare de tip discret, cu tabloul de repartiţie

$$\begin{pmatrix} x_i \\ p_i \end{pmatrix} ; \ i = 1, 2, \ldots$$

Variabila aleatoare X^k va avea atunci tabloul de repartiţie:

$$\begin{pmatrix} x_i^k \\ p_i \end{pmatrix} ; \ i = 1, 2, \ldots$$

şi deci, pe baza definiţiei mediei unei variabile aleatore de tip discret, putem scrie:

$$M(X^k) = \sum_i x_i^k p_i \ , \ K \in \mathbb{N}$$

ceea ce justifică relaţia (7.6.7).

b) Fie acum X o variabilă aleatoare de tip continuu, cu funcţia de densitate $f, f: \mathbb{R} \to \mathbb{R}$.

Atunci funcţia de densitate φ, a variabilei aleatoare X^k, este dată (vezi exerciţiul 12 de la sfîrşitul paragrafului precedent) de:

$$\varphi(x) = \begin{cases} 0 & , \ x \le 0 \\ \dfrac{1}{2p} x^{\frac{1}{2p}-1} \left[f(x^{\frac{1}{2p}}) + f(-x^{\frac{1}{2p}}) \right] & ; \ x > 0 \end{cases}$$

în cazul în care $k = 2p$ şi de:

$$\varphi(x) = \frac{1}{2p+1} x^{\frac{-2p}{2p+1}} f(x^{\frac{1}{2p+1}}) \ , \ x \in \mathbb{R}$$

în cazul în care $k = 2p + 1$.

Avînd în vedere acum definiţia mediei unei variabile aleatoare de tip continuu rezultă:

$$M(X^k) = \int_0^{\infty} x \frac{1}{k} x^{\frac{1}{k}-1} \left[f(x^{\frac{1}{k}}) + f(-x^{\frac{1}{k}}) \right] dx \qquad 7.6.8a$$

în cazul în care k este un număr par, $k = 2p$, şi:

$$M(X^k) = \int_{-\infty}^{+\infty} x \frac{1}{k} x^{\frac{1-k}{k}} f(x^{\frac{1}{k}}) dx \qquad 7.6.8b$$

în cazul în care k este un număr impar, $k = 2p + 1$.

Dacă în integrala din relaţiile (7.6.8 a,b) se face substituţia $X^{\frac{1}{k}} = t$, obţinem relaţiile echivalente:

$$M(X^k) = \int_0^{\infty} t^k [f(t) + f(-t)] dt \quad pentru \quad k = 2p$$

$$M(X^k) = \int_{-\infty}^{+\infty} t^k f(t) dt \quad pentru \quad k = 2p + 1$$

Dar, atunci cînd $k = 2p$, are loc:

$$\int_0^{\infty} t^k f(-t) dt = \int_{-\infty}^{0} t^k f(t) dt$$

şi deci putem scrie:

$$\int_0^{\infty} t^k [f(t) + f(-t)] dt = \int_{-\infty}^{+\infty} t^k f(t) dt$$

ceea ce arată că atît în cazul $k = 2p$, cît şi în cazul cînd $k = 2p + 1$ are loc relaţia:

$$M(X^k) = \int_{-\infty}^{+\infty} t^k f(t)\, dt$$

şi, prin urmare, cu aceasta am dovedit şi relaţia (7.6.7').

BCU "ASICENTAL UNIVERSITY LIBRARY

Definiţia 7.6.5.

Fie X o variabilă aleatoare definită peste un cîmp de probabilitate (Ω, K, P). Se numeşte moment centrat de ordin k, $k: \mathbb{N}$, al acestei variabile aleatoare momentul iniţial de ordin k al variabilei aleatoare $Y = X - M(X)$, (variabila aleatoare abatere a variabilei X de la media sa).

Momentul centrat de ordin k al variabilei aleatoare X îl vom nota prin $\mu_k(X)$. Conform definiţiei, urmează:

$$\mu_k(X) = M[(X - M(X))^k]$$

Se observă imediat că momentul centrat de ordin al doilea pentru o variabilă aleatoare coincide cu dispersia variabilei aleatoare respective, adică $\mu_2(X) = D(X)$. Putem dovedi că între momentele iniţiale şi momentele centrate ale unei variabile aleatoare X au loc următoarele relaţii de legătură:

$$\mu_k(X) = \sum_{i=0}^{k} (-1)^i C_k^i \alpha_{k-i}(X)\, \alpha_1^i(X) \; ; \quad k = 1, 2, \ldots \qquad 7.6.9$$

$$\alpha_k(X) = \sum_{i=0}^{k} C_k^i \mu_{k-i}(X)\, \alpha_1^i(X) \; ; \quad k = 1, 2, \ldots$$

Înainte de a justifica aceste relaţii, să notăm faptul că prima dintre relaţiile (7.6.9) arată că pentru a determina momentul centrat de ordin k, $k \geq 1$, al unei variabile aleatoare este suficient să se cunoască momentele iniţiale ale acestei variabile aleatoare, pînă la ordinul k inclusiv.

Pentru a dovedi relaţiile (7.6.9), să observăm că, ţinînd cont de binomul lui Newton, putem scrie:

$$(X - M(X))^k = \sum_{i=0}^{k} C_k^i X^{k-i} (-M(X))^i \; ; \; k = 1, 2, \ldots \qquad 7.6.10$$

$$X^k = (X - M(X) + M(X))^k = \sum_{i=0}^{k} C_k^i [X - M(X)]^{k-i} [M(X)]^i \; ; \; k = 1, 2, \ldots$$

Deoarece $M(X) = \alpha_1(X)$, $M(X^{k-i}) = \alpha_{k-i}(X)$, $M(X - M(X))^{k-i} = \mu_{k-i}(X)$, pe baza proprietăţilor mediei unei variabile aleatoare, din (7.6.10) rezultă imediat relaţiile (7.6.9). Deci relaţiile (7.6.9) sînt demonstrate.

Inegalitatea lui Cebîşev

Vom demonstra acum o inegalitate, cunoscută sub numele de inegalitatea lui Cebîşev, care se va dovedi deosebit de utilă în partea de statistică matematică.

Teorema 7.6.1.

Dacă X este o variabilă aleatoare definită peste un cîmp de probabilitate $(\Omega, K, \mathcal{P})$ şi dacă X are media $M(X)$ şi dispersia $D(X)$, atunci are loc inegalitatea:

$$P(|X - M(X)| < \varepsilon) \geq 1 - \frac{D(X)}{\varepsilon^2}, \ (\forall)\, \varepsilon > 0 \qquad 7.6.11$$

Demonstraţie: Vom demonstra această teoremă numai pentru cazul variabilelor aleatoare de tip discret sau de tip continuu.

a) Să presupunem că X este o variabilă aleatoare de tip continuu cu funcţia de densitate $f, f: \mathbf{R} \to \mathbf{R}$. Dacă notăm $M(X) = m$, atunci, pe baza relaţiei (7.6.6), putem scrie:

$$D(X) = \int_{-\infty}^{+\infty} (x - m)^2 f(x)\, dx \qquad 7.6.12$$

Pe de altă parte, deoarece funcţia de repartiţie a unei variabile aleatoare de tip continuu este continuă urmează:

$$P(|X - m| < \varepsilon) = P(m - \varepsilon < X < m + \varepsilon) = F(m + \varepsilon) - F(m - \varepsilon) \qquad 7.6.13$$

unde F este funcţia de repartiţie a variabilei aleatoare X. Dar, pe baza proprietăţilor integralei definite, putem scrie:

$$\int_{-\infty}^{+\infty} (x - m)^2 f(x)\, dx = \int_{-\infty}^{m-\varepsilon} (x - m)^2 f(x)\, dx + \int_{m-\varepsilon}^{m+\varepsilon} (x - m)^2 f(x)\, dx +$$

$$+ \int_{m+\varepsilon}^{+\infty} (x - m)^2 f(x)\, dx \geq \int_{-\infty}^{m-\varepsilon} (x - m)^2 f(x)\, dx + \int_{m+\varepsilon}^{+\infty} (x - m)^2 f(x)\, dt$$

deoarece $(x - m)^2 f(x) \geq 0$, $(\forall) x \in \mathbf{R}$.

Totodată, din faptul că $f(x) \geq 0$, $(\forall) x \in \mathbb{R}$ şi $(x-m)^2 \geq \varepsilon^2$ pentru $x \in (-\infty, m-\varepsilon) \cup [m+\varepsilon, +\infty]$, rezultă:

$$\int_{-\infty}^{m-\varepsilon} (x-m)^2 f(x)\,dx \geq \varepsilon^2 \int_{-\infty}^{m-\varepsilon} f(x)\,dx = \varepsilon^2 F(m-\varepsilon)$$

$$\int_{m+\varepsilon}^{-\infty} (x-m)^2 f(x)\,dx \geq \varepsilon^2 \int_{m+\varepsilon}^{+\infty} f(x) = \varepsilon^2 (1 - F(m+\varepsilon))$$

şi deci:

$$\int_{-\infty}^{+\infty} (x-m)^2 f(x)\,dx \geq \varepsilon^2 (1 - [F(m+\varepsilon)) - F(m-\varepsilon)])$$

Avînd în vedere relaţiile (7.6.12), (7.6.13), constatăm că are loc relaţia

$$D(X) \geq \varepsilon^2 [1 - P(|X-m| < \varepsilon)]$$

sau echivalent:

$$P(|X-m| < \varepsilon) \geq 1 - \frac{D(X)}{\varepsilon^2}$$

ceea ce dovedeşte inegalitatea (7.6.11) pentru variabile aleatoare de tip continuu.

b) Fie acum X o variabilă aleatoare de tip discret avînd tabloul de repartiţie:

$$\begin{pmatrix} x_i \\ p_i \end{pmatrix}, \; i = 1, 2, \ldots$$

Deoarece m notează media acestei variabile aleatoare, pe baza relaţiei (7.6.5), putem scrie:

$$D(X) = \sum_i (x_i - m)^2 p_i \qquad 7.6.12'$$

Fie acum mulţimile:

$$I_1 = \{ i ; x_i - m \geq \varepsilon \}$$

$$I_2 = \{ i ; x_i - m \leq \varepsilon \}$$

$$I_3 = \{ i ; -\varepsilon < x_i - m < \varepsilon \}$$

Este atunci clar că au loc relaţiile:

$$\sum_i (x_i - m)^2 p_i \geq \sum_{i \in I_1} (x_i - m)^2 p_i + \sum_{i \in I_2} (x_i - m)^2 p_i \geq$$

$$\geq \varepsilon^2 \left(\sum_{i \in I_1} p_i + \sum_{i \in I_2} p_i \right) = \varepsilon^2 \left(1 - \sum_{i \in I_3} p_i \right)$$

Dar, din modul în care a fost definită mulţimea I_3, rezultă:

$$\sum_{i \in I_3} p_i = P(|X - m| < \varepsilon) \qquad 7.6.13'$$

şi atunci, avînd în vedere şi relaţia (7.6.12'), urmează:

$$D(X) \geq \varepsilon^2 (1 - \mathcal{P}(|X - m| < \varepsilon))$$

sau echivalent:

$$P(|X - m| < \varepsilon) \geq 1 - \frac{D(X)}{\varepsilon^2}$$

ceea ce dovedeşte că inegalitatea (7.6.11) este adevărată şi în cazul variabilelor aleatoare de tip discret.

Exerciţii

1. Să se arate că dispersia unei variabile aleatoare de tip discret este egală cu zero atunci şi numai atunci cînd variabila aleatoare respectivă este o constantă. Este adevărată această afirmaţie şi în cazul variabilelor aleatoare de tip continuu?

2. Fie X o variabilă aleatoare de tip continuu cu funcţia de densitate $f, f: \mathbb{R} \to \mathbb{R}$. Să se arate că variabila aleatoare normată, asociată văriabilei aleatoare X, are functia de densitate φ, $\varphi : \mathbb{R} \to \mathbb{R}$, definită prin:

$$\varphi(x) = \sqrt{D(X)} \cdot f(x\sqrt{D(X)} + m), \quad x \in \mathbb{R}$$

3. Se consideră o variabilă aleatoare X pentru care momentele iniţiale $\alpha_k(X), k = 1, 2, 3$, sînt date prin:

$$\alpha_k(X) = \frac{1}{k+1}, \; k = 1, 2, 3$$

Să se determine dispersia şi momentul centrat de ordinul 3 pentru variabila aleatoare X.

4. Se consideră variabila aleatoare de tip discret, X, cu tabloul de repartiţie:

$$\begin{pmatrix} 0 & 1 & 2 & 3 \\ \left(\frac{2}{3}\right)^3 & \left(\frac{2}{3}\right)^2 & \frac{2}{3^2} & \frac{1}{3^3} \end{pmatrix}$$

Să sedetermine media şi dispersia variabilei aleatoare X.

5. Fie X o variabilă aleatoare de tip continuu a cărei funcţie de densitate este $f, f: \mathbf{R} \to \mathbf{R}$, definită prin:

$$f(x) = \begin{cases} 0 & ; x < 0 \\ e^{-x} & ; x \geq 0 \end{cases}$$

Să se determine media şi dispersia variabilei aleatoare X.

6. Fie X o variabilă aleatoare a cărei funcţie de repartiţie F, $F: \mathbf{R} \to \mathbf{R}$, este dată prin:

$$F(x) = \begin{cases} 0 & ; x < 0 \\ 1 - e^{-\lambda x} & ; x \geq 0 \end{cases}$$

unde λ este un parametru real pozitiv.

Să se determine momentele iniţiale şi centrate ale variabilei aleatoare X, pînă la ordinul 3 inclusiv.

Corelaţie. Coeficient de corelaţie

În practică o variabiliă aleatoare unidimensională corespunde unei caracteristici numerice a unei colectivităţi. Dacă la o anumită colectivitate ne interesează mai multe caracteristici numerice, atunci, în funcţie de acestea, vom putea asocia colectivităţi mai multe variabile aleatoare. Desigur nu întotdeauna astfel de caracteristici sînt independente şi deci nici variabilele corespunzătoare nu sînt

întotdeauna independente. Vorbind în general, dacă se consideră mai multe variabile aleatoare definite peste un cîmp de probabilitate poate apare necesar să se stabilească: 1. dacă aceste variabile aleatoare sînt independente; 2. dacă nu sînt independente, care este legătura dintre aceste variabile. În cele ce urmează vom considera cazul cînd este nevoie să se răspunde la astfel de întrebări doar pentru două variabile aleatoare X, Y, definite peste un cîmp de probabilitate dat.

Faptul dacă variabilele X şi Y sînt independente poate fi stabilit, cel puţin în principiu, verificînd îndeplinirea condiţiei de independenţă a acestora. În continuare vom vedea cum se poate verifica independenţa a două variabile aleatoare corespunzătoare unor caracteristici numerice ale unei colectivităţi.

Remarcăm faptul că atunci cînd între două caracteristici ale unei colectivităţi există o relaţie de interdependenţă se spune că aceste caracteristici sînt *corelate* (de exemplu vîrsta şi greutatea indivizilor dintr-o colectivitate umană). Subliniem faptul foarte important că atunci cînd se spune că două astfel de caracteristici sînt corelate, nu înseamnă că valoarea uneia dintre ele determină în mod unic valoarea celeilalte. Revenind la cazul general, se ştie că atunci cînd două variabile aleatoare X, Y sînt independente are loc:

$$D(X+Y) = D(X) + D(Y)$$

iar dacă X şi Y sînt două variabile aleatoare oarecare:

$$D(X+Y) = D(X) + D(Y) + 2M[(X - M(X))(Y - M(Y))]$$

Putem considera atunci că numărul $M(X - M(X))(Y - M(Y))$ exprimă, într-o anumită măsură, gradul de interdependenţă a variabilelor aleatoare X şi Y. Dacă variabilele aleatoare X şi Y sînt independente atunci acest număr este egal cu zero. Este necesar însă să subliniem că atunci cînd acest număr este zero nu rezultă obligatoriu faptul că variabilele aleatoare X, Y sînt independente. Pe de altă parte este sigur că atunci cînd acest număr, $M[(X - M(X))(X - M(Y))]$, este diferit de zero variabilele aleatoare X şi Y nu sînt independente.

Definiţia 7.6.6.

Fie X şi Y două variabile aleatoare definite peste un cîmp de probabilitate $(\Omega, K, \mathcal{P})$. Prin *corelaţie* sau *covarianţă* a acestor variabile aleatoare înţelegem numărul notat $cov(X, Y)$ şi definit prin:

$$cov(X, Y) = M[(X - M(X)(Y - M(Y))]$$

În cazul cînd variabilele aleatoare şi Y sînt astfel încît $cov(X, Y) \neq 0$ spunem că variabilele X, Y sînt corelate, iar dacă $cov(X, Y) \neq 0$ vom spune că variabilele aleatoare X, Y nu sînt corelate. Din definiţia corelaţiei a două variabile aleatoare şi pe baza proprietăţilor mediei unei variabile aleatoare rezultă imediat că are loc:

$$cov(X, Y) = M(X.Y) - M(X)M(Y)$$

Definiţia 7.6.7.

Se numeşte coeficient de corelaţie a două variabile aleatoare X, Y, numărul $\rho(X, Y)$ definit prin:

$$\rho(X, Y) = \frac{cov(X, Y)}{\sigma_X \cdot \sigma_Y}$$

unde prin $\sigma_X, \sigma_Y, (\sigma_X, \sigma_Y \neq 0)$, am notat abaterile medii pătratice ale variabilelor X, Y, respectiv.

Observaţia 1. Dacă $\sigma_X = 0$, $(\sigma_Y = 0)$, atunci variabila aleatoare $X, (Y)$, poate fi considerată constantă, ceea ce implică $cov(X, Y) = 0$ şi atunci prin definiţie considerăm $\rho(X, Y) = 0$.

Observaţia 2. Putem constata imediat că dacă se calculează corelaţia variabilelor normate asociate variabilelor X, Y atunci acesta este coeficientul de corelaţie a variabilelor X, Y, adică:

$$\rho(X, Y) = cov\left(\frac{X - M(x)}{\sigma_X}, \frac{Y - M(Y)}{\sigma_Y}\right)$$

Teorema 7.6.2.

Dacă X, Y sînt două variabile aleatoare definite acelaşi cîmp de probabilitate atunci are loc relaţia:

$$-1 \leq \rho(X, Y) \leq 1$$

Demonstraţie: Să notăm cu $\bar{X}, \bar{Y}$ variabilele aleatoare normate asociate variabilelor aleatoare X, Y, respectiv.

Atunci $M(\bar{X}) = 0$, $M(\bar{Y}) = 0$, $D(\bar{X}) = 1 = M(\bar{X}^2)$, $D(\bar{Y}) = 1 = M(\bar{Y}^2)$ şi, în plus, $cov(\bar{X}, \bar{Y}) = \rho(X, Y)$.

Să observăm acum că pentru oricare $t \in \mathbf{R}$ are loc:

$$M[(\bar{X} + t\bar{Y})^2] \geq 0$$

de unde rezultă:

$$M(\bar{X}^2) + 2tM(\bar{X}\bar{Y}) + t^2 M(\bar{Y}^2) \geq 0\ ,\ (\forall) t \in \mathbf{R}$$

şi cum $M(\bar{X}^2) = 1, M(\bar{Y}^2) = 1$, urmează:

$$t^2 + 2t\, cov(\bar{X}, \bar{Y}) + 2 \geq 0; (\forall) t \mathbf{R}$$

Deoarece $cov(\bar{X}, \bar{Y}) = \rho(X, Y)$, din această relaţie urmează evident:

$$\rho^2(X, Y) \leq 1$$

ceea ce demonstrează teorema:

Teorema 7.6.3.

Dacă variabilele aleatoare X, Y sînt astfel încît $\rho^2(X, Y) = 1$ atunci există două constante reale $a, b, a \neq 0$ astfel încît $X = aY + b$, adică între variabilele aleatoare X şi Y există o relaţie de dependenţă liniară.

Demonstraţie: După cum am văzut în demonstraţia teoremei precedente, pentru orice $t \in \mathbf{R}$ are loc relaţia:

$$M[(\bar{X} + t\bar{Y})^2] = t^2 + 2t\rho(X, Y) + 1$$

unde $\bar{X}, \bar{Y}$ sînt variabilele aleatoare normate asociate variabilelor X, Y, respectiv.

Relaţia fiind adevărată şi pentru $t = -\rho(X, Y)$, urmează:

$$M[(\bar{X} - \rho(X, Y)\bar{Y})^2] = 1 - \rho^2(X, Y)$$

şi, deoarece prin ipoteză se presupune că $\rho^2(X, Y) = 1$, aceasta relaţie implică:

$$M[(\bar{X} - \rho(X,Y)\bar{Y})^2] = 0$$

Avînd în vedere că variabila aleatoare $(\bar{X} - \rho(X,Y)\bar{Y})^2$ are toate valorile nenegative, media sa este egală cu zero numai toate valorilor nenegative, media sa este egală cu zero numai dacă:

$$\bar{X} - \rho(X,Y)\bar{Y} = 0 \qquad 7.6.14$$

Deoarece $\bar{X} = \frac{X - M(X)}{\sigma_x}$, $\bar{Y} = \frac{Y - M(Y)}{\sigma_Y}$, din relaţia (7.6.14) urmează:

$$X = aY + b$$

în care:

$$a = \frac{\sigma_X}{\sigma_Y}\rho(X,Y), \quad a \neq 0$$

$$b = M(X) - M(Y)\frac{\sigma_X}{\sigma_Y}\rho(X,Y)$$

ceea ce demonstrează teorema.

Se poate dovedi fără dificultate că este adevărată şi reciproca acestei teoreme, adică atunci cînd între două variabile X, Y există o relaţie de dependenţă de forma $X = aY + b$, $a \neq 0$ atunci:

$$\rho^2(X,Y) = 1$$

Să presupunem acum că două variabile aleatoare X, Y sînt corelate, adică $\rho(X,Y) \neq 0$. După cum am văzut, în cazul cînd $\rho(X,Y) = \pm 1$, există două numere reale $a, b, a \neq 0$, astfel încît $X = aY + b$, iar în cazul cînd $\rho(X,Y) \neq \pm 1$ între variabilele aleatoare X, Y nu există relaţii de dependenţă liniară. Să observăm că existenţa unei relaţii de forma $X = aY + b$ între variabilele aleatoare X, Y este echivalentă cu:

$$M(X - aY - b) = 0, \quad D(X - aY - b) = 0$$

iar existenţa unei relaţii de forma $Y = aX + b$ este echivalentă cu:

$$M(Y - aX - b) = 0, \quad D(Y - aX - b) = 0$$

Dacă două variabile aleatoare X, Y sînt corelate ne propunem să

determinăm două numere a_1, b_1 astfel încît să fie îndeplinite condiţiile:

$$M(Y - a_1X - b_1) = 0$$
$$D(Y - a_1X - b_1) \min$$

Se constată imediat că:

$$D(Y - a_1X - b_1) = D(Y) + a_1^2 D(X) - 2a_1 cov(X, Y) \qquad 7.6.15$$

şi atunci $D(Y - a_1X - b_1)$ este minimă atunci cînd:

$$a_1 = \frac{cov(X, Y)}{D(X)} = \rho(X, Y)\frac{\sigma_Y}{\sigma_X}$$

Pe de altă parte, dacă a_1, b_1 trebuie astfel determinaţi încît $M(Y) = a_1 M(X) + b_1$ urmează:

$$b_1 = M(Y) - a_1 M(X)$$

şi, ţinînd cont de valoarea lui a_1 determinată mai sus, rezultă:

$$b_1 = M(Y) - M(X)\frac{\sigma_Y}{\sigma_X}\rho(X, Y) \qquad 7.6.16$$

Dacă vom căuta două numere reale a_2, b_2 astfel încît

$$M(X - a_2Y - b_2) = 0$$
$$D(X - a_2Y - b_2) \min$$

atunci, procedînd în acelaşi mod, vom obţine:

$$a_2 = \frac{\sigma_X}{\sigma_Y}\rho(X, Y)$$
$$b_2 = M(X) - M(Y)\frac{\sigma_X}{\sigma_Y}\rho(X, Y) \qquad 7.6.17$$

Definiţia 7.6.8.

Dacă pentru două variabile aleatoare definite peste un acelaşi cîmp probabilitate coeficientul de corelaţie este nenul şi dacă (x, y) sînt coordonatele unui punct din plan, atunci dreapta de ecuaţie:

$$y = a_1 x + b_1$$

se numeşte dreapta de regresie a variabilei aleatoare Y în raport cu variabila aleatoare X, iar dreapta de ecuaţie:

$$x = a_2 y + b_2$$

se numeşte dreapta de regresie a variabilei aleatoare X în raport cu variabila aleatoare Y, coeficienţii a_1, a_2, b_1, b_2 pentru aceste drepte fiind daţi prin relaţiile (7.6.15)-(7.6.17).

Asupra semnificaţiei dreptelor de regresie a două variabile aleatoare asociate la două caracteristici ale unei colectivităţi vom reveni în partea de statistică matematică.

§ 7. REPARTIŢII PROBABILISTICE DISCRETE UZUALE

În acest capitol vom prezenta cîteva repartiţii probabilistice, de tip discret şi de tip continuu, utilizate în mod frecvent în rezolvarea problemelor practice cu mijloacele statistice matematice. Pentru fiecare dintre aceste repartiţii vom calcula media şi dispersia, evidenţiind atunci cînd este necesar şi alte caracteristici ale acestor repartiţii.

1. Repartiţia hipergeometrică

Peste cîmpul de probabilitate asociat unei experienţe organizate conform schemei bilei nerevenite se poate construi în mod natural o variabilă aleatoare discretă ale cărei valori reprezintă numărul de bile albe extrase din urnă. Variabilele aleatoare discrete cu tabloul de repartiţie de forma tabloului de repartiţie pentru o astfel de variabilă aleatoare se grupează într-o clasă, repartiţia unei variabile aleatoare din această clasă purtînd denumirea de repartiţie hipergeometrică.

Definiţia 7.7.1.

Fie (Ω, K, P) un cîmp de probabilitate şi fie X o variabilă aleatoare dis-

cretă definită peste acest cîmp de probabilitate. Vom spune că variabila aleatoare X are o repartiţie hipergeometrică dacă tabloul său de repartiţie este de forma:

$$\begin{pmatrix} 0 & 1 & 2 \dots n \\ p_0 & p_1 & p_2 \cdots p_n \end{pmatrix} \qquad 7.7.1$$

în care $p_k = \dfrac{C_a^k C_b^{n-k}}{C_{a+b}^n}$, $k = 0, 1, \dots, n$, a şi b fiind două numere naturale fixate, $a + b \geq n$.

Să dovedim că tabloul (7.7.1) este într-adevăr tabloul de repartiţie al unei variabile aleatoare, adică:

a) $p_k \geq 0$, $k = 0, 1, \dots, n$

b) $\sum_{k=0}^{n} p_k = 1$

Prima dintre aceste condiţii este evident îndeplinită deoarece $C_{a+b}^n > 0$; $C_a^k, C_b^{n-k} \geq 0$; $k = 0, 1, \dots, n$. Pentru a dovedi că este îndeplinită şi condiţia b), ne vom folosi de identitatea:

$$(1 + x)^a (1 + x)^b = (1 + x)^{a+b} \ , \ (\forall) x \in \mathbb{R} \qquad 7.7.2$$

Considerînd $n \leq a + b$ constatăm că în primul membru al identităţii (7.7.2) coeficientului lui x^n este:

$$\sum_{k=0}^{n} C_a^k C_b^{n-k}$$

iar în cel de al doilea membru al acestei identităţi coeficientul lui x^n este C_{a+b}^n. Prin urmare, din identitatea (7.7.2) urmează:

$$\sum_{k=0}^{n} C_a^k C_b^{n-k} = C_{a+b}^n \qquad 7.7.3$$

ceea ce este echivalent cu $\dfrac{1}{C_{a+b}^n} \sum_{k=0}^{n} C_a^k C_b^{n-k} = 1$.

BCU IAŞI CENTRAL UNIVERSITY LIBRARY

Deoarece, din definiţia probabilităţilor $p_k, k = 0, 1, \ldots, n$ rezultă:

$$\sum_{k=0}^{n} p_k = \frac{1}{C_{a+b}^{n}} \sum_{k=0}^{n} C_a^k C_b^{n-k}$$

din cele arătate mai sus urmează că $\sum_{k=0}^{n} p_k = 1$, ceea ce dovedeşte că tabloul (7.7.1) este tabloul de repartiţie al unei variabile aleatoare.

Media unei variabile aleatoare cu repartiţie hipergeometrică

Din tabloul de repartiţie al unei variabile aleatoare X, cu repartiţie hipergeometrică, rezultă că media acesteia este:

$$M(X) = \sum_{k=0}^{n} k \frac{C_a^k C_b^{n-k}}{C_{a+b}^{n}} = \frac{1}{C_{a+b}^{n}} \sum_{k=0}^{n} k C_a^k C_b^{n-k} \qquad 7.7.4$$

Pentru determinarea unei forme simple a mediei acestei variabile aleatoare vom calcula $\sum_{k=0}^{n} k C_a^k C_b^{n-k}$. Dar:

$$\sum_{k=0}^{n} k C_a^k C_b^{n-k} = \sum_{k=1}^{n} k \frac{a!}{k!\,(a-k)!} C_b^{n-k} = \qquad 7.7.5$$

$$= \sum_{k=1}^{n} \frac{a!}{(k-1)!\,(a-k)!} C_b^{n-k} = a \sum_{k=1}^{n} C_{a-1}^{k-1} C_b^{n-k} = a \sum_{h=0}^{n-1} C_{a-1}^{h} C_b^{n-1-h}$$

Deoarece relaţia (7.7.3) este adevărată pentru orice a, b, n urmează că $\sum_{h=0}^{n-1} C_{a-1}^{h} C_b^{n-1-h} = C_{a+b-1}^{n-1}$ şi atunci, din relaţiile (7.7.5) rezultă:

$$\sum_{k=0}^{n} k C_a^k C_b^{n-k} = a C_{a+b-1}^{n-1} \qquad 7.7.6$$

Deci putem scrie:

$$M(X) = \frac{1}{C_{a+b}^{n}} a\, C_{a+b-1}^{n-1} \qquad 7.7.7$$

de unde, după efectuarea simplificărilor, se obţine:

$$M(X) = n \frac{a}{a+b} \qquad 7.7.8$$

Dispersia unei variabile aleatoare cu repartiţie hipergeometrică

După cum se ştie, pentru orice variabilă aleatoare X are loc relaţia:

$$D(X) = M(X)^2 - [M(X)]^2 \qquad 7.7.9$$

şi deci, pentru a determina dispersia unei variabile aleatoare cu repariţie hipergeometrică, este suficient să determinăm $M(X^2)$. Din tabloul de repartiţie al unei variabile aleatoare cu repartiţie hipergeometrică rezultă:

$$M(X^2) = \sum_{k=0}^{n} k^2 \frac{C_a^k C_b^{n-k}}{C_{a+b}^{n}} = \frac{1}{C_{a+b}^{n}} \sum_{k=1}^{n} k^2 C_a^k C_b^{n-k} \qquad 7.7.10$$

Deoarece $k^2 = k(k-1) + k$, putem scrie succesiv relaţiile:

$$\sum_{k=1}^{n} k^2 C_a^k C_b^{n-k} = \sum_{k=1} k(k-1) C_a^k C_b^{n-k} + \sum_{k=1}^{n} k C_a^k C_b^{n-k} =$$

$$= \sum_{k=2}^{n} a(a-1) C_{a-2}^{k-2} C_b^{n-k} + n \frac{a}{a+b} C_{a+b}^{n} = \qquad 7.7.11$$

$$= a(a-1) \sum_{h=0}^{n-2} C_{a-2}^{n-2} C_b^{n-2-h} + n \frac{a}{a+b} C_{a+b}^{n}$$

Utilizînd acum relaţia (7.7.3) în care în loc de a se pune $a-2$ şi în loc de n punem $n-2$, rezultă:

$$\sum_{h=o}^{n-2} C_{a-2}^{h} C_b^{n-2-h} = C_{a+b-2}^{n-2} \qquad 7.7.12$$

Din (7.7.11) şi (7.7.12) urmează atunci:

$$\sum_{k=1}^{n} k^2 C_a^k C_b^{n-k} = a(a-1) C_{a+b-2}^{n-2} + n \frac{a}{a+b} C_{a+b}^{n} \qquad 7.7.13$$

Prin urmare are loc relaţia:

$$M(X^2) = a(a-1) \frac{C_{a+b-2}^{n-2}}{C_{a+b}^{n}} + n \frac{a}{a+b} \qquad 7.7.14$$

şi, după efectuarea simplificărilor, obţinem:

$$M(X^2) = \frac{na}{a+b}\left[\frac{(a-1)(n-1)}{a+b-1} + 1\right] \qquad 7.7.15$$

Deci putem scrie:

$$M(X^2) = \frac{na}{a+b}\left[\frac{(a-1)(n-1)}{a+b-1} + 1\right] - \left(\frac{na}{a+b}\right)^2$$

de unde, după efectuarea calculelor, obţinem:

$$D(X) = n \frac{a}{a+b} \cdot \frac{b}{a+b} \cdot \frac{a+b-n}{a+b-1} \qquad 7.7.16$$

Să observăm că, dacă pentru o variabilă aleatoare cu repartiţie hipergeometrică notăm $p = \frac{a}{a+b}$, $q = \frac{b}{a+b}$, $t = a + b$ atunci:

$$\begin{aligned} M(X) &= np \\ D(X) &= npq \frac{t-n}{t-1} \end{aligned} \qquad 7.7.17$$

şi, dacă interpretăm o astfel de variabilă aleatoare ca fiind asociată unei experienţe organizate conform schemei bilei nerevenite, experienţa constînd în extragerea succesivă a n bile dintr-o urnă în care sînt $\underline{a}$ bile albe şi $\underline{b}$ bile negre, atunci p este probabilitatea de a extrage o bilă albă din urnă (proporţia de bile albe din urnă), q este probabilitatea de a extrage o bilă neagră din urnă, $q = 1 - p$, iar t este numărul total de bile din urnă.

2. Repartiţia binomială

Definiţia 7.7.2.

Fie (Ω, K, P) un cîmp de probabilitate şi X o variabilă aleatoare discretă definită peste acest cîmp de probabilitate. Vom spune că variabila aleatoare X are o repartiţie binomială dacă tabloul său de repartiţie este de forma:

$$\begin{pmatrix} 0 & 1 & 2 \dots n \\ p_0 & p_1 & p_2 \cdots p_n \end{pmatrix} \qquad 7.7.18$$

unde $p_k = C_n^k p^k q^{n-k}$, iar $p + q = 1$, $p > 0$, $q > 0$.

Să arătăm întîi că tabloul (7.7.18) este tabloul de repartiţie al unei variabile aleatoare discrete.

Într-adevăr, deoarece $p, q > 0$, $C_n^k \geq 0$, $k = 0, 1, \dots, n$ rezultă că $p_k \geq 0$, $k = 0, 1, 2, \dots, n$.

Pe de altă parte, deoarece $(p + q)^n = 1$ urmează:

$$\sum_{k=0}^{n} C_n^k p^k q^{n-k} = (p + q)^n = 1$$

adică $\sum_{k=1}^{n} p_k = 1$, ceea ce dovedeşte că tabelul (7.7.18) este tabloul de repartiţie al unei variabile aleatoare discrete.

Înainte de a trece la calculul mediei şi dispersiei unei astfel de variabile aleatoare vom demonstra o teoremă, care se va dovedi utilă pentru determinarea rapidă a acestor caracteristici numerice.

Teorema 7.7.1.

Fie X_i, $i = 1, \dots, n$, n variabile aleatoare discrete independente definite peste acelaşi cîmp de probabilitate, toate avînd acelaşi tablou de repartiţie :

$$\begin{pmatrix} 0 & 1 \\ q & p \end{pmatrix}$$

în care $q, p > 0$, $p + q = 1$. Atunci variabila aleatoare $X = \sum_{i=1}^{n} X_i$ are o repartiţie binomială.

BIJLIOTECA CENTRALĂ UNIVERSITARĂ LIBRARY

Demonstraţie: Este clar că variabila aleatoare $X = \sum_{i=1}^{n} X_i$ este o variabilă aleatoare discretă cu valori în mulţimea $\{0, 1, 2, \ldots, n\}$, variabila aleatoare X luînd valoarea k, $0 \le k \le n$, atunci cînd exact k dintre variabilele aleatoare X_i iau valoarea 1, iar celelalte iau valoarea zero, adică $X(\omega) = k$ dacă $X_i(\omega) = 1$, $i \in \{i_1, \ldots, i_k\}$, şi $X_i(\omega) = 0$ pentru $i \notin \{i_1, \ldots, i_k\}$. Deoarece variabilele aleatoare X_i sînt independente urmează:

$$P(X_{i_1})(\omega) = 1, \ldots, X_{i_k}(\omega) = 1, X_i(\omega) = 0, i \notin \{i_1, \ldots, i_k\} = p^k q^{n-k}$$

şi atunci $P(X(\omega) = k) = C_n^k p^k q^{n-k}$, din cauza faptului că există C_n^k moduri de alegere a k variabile aleatoare X_j, $i \in \{1, 2, \ldots, n\}$, care să ia valoarea 1, iar celelalte $(n - k)$ să ia valoarea zero. Deci tabloul de repartiţie al variabilei aleatoare $X = \sum_{i=1}^{n} X_i$ este:

$$\begin{pmatrix} k \\ C_n^k p^k q^{n-k} \end{pmatrix}; \; k = 0, 1, \ldots, n$$

ceea ce arată că X are o repartiţie binomială şi teorema este demonstrată.

Media unei variabile aleatoare cu repartiţie binomială

Pe baza teoremei precedente putem spune că dacă X este o variabilă aleatoare cu repartiţie binomială, atunci considerînd variabilele aleatoare X_i, $i = 1, \ldots, n$, independente, avînd toate tablourile de repartiţie:

$$\begin{pmatrix} 0 & 1 \\ q & p \end{pmatrix}$$

putem scrie $X = \sum_{i=1}^{n} X_i$.

Din proprietăţile mediei unei variabile aleatoare rezultă atunci că are loc:

$$M(X) = \sum_{i=1}^{n} M(X_i) \qquad 7.7.19$$

Dar din tabloul de repartiţie al variabilelor aleatoare X_i, $i = 1,\ldots,n$, este evident că $M(X_i) = p$, $i = 1,\ldots,n$ şi atunci:

$$M(X) = \sum_{i=1}^{n} p \qquad 7.7.20$$

ceea ce înseamnă:

$$M(X) = np \qquad 7.7.21$$

Dispersia unei variabile aleatoare cu repartiţie binomială

Ţinînd iarăşi cont de faptul că o variabilă aleatoare X, cu repartiţie binomială, se poate pune sub forma $X = \sum_{i=1}^{n} X_i$ în care X_i, $i = 1,\ldots,n$, sînt variabile aleatoare independente avînd acelaşi tablou de repartiţie:

$$\begin{pmatrix} 0 & 1 \\ q & p \end{pmatrix}$$

din proprietăţile dispersiei unei variabile aleatoare rezultă:

$$D(X) = \sum_{i=1}^{n} D(X_i) \qquad 7.7.22$$

Deoarece $D(X_i) = M(X_i^2) - [M(X_i)]^2 = p - p^2 = pq$, $i = 1,\ldots,n$ din (7.7.22) rezultă nemijlocit:

$$D(X) = npq \qquad 7.7.23$$

Remarcăm faptul că o variabilă aleatoare cu repatiţie binomială poate fi considerată ca fiind asociată unei experienţe organizate conform schemei bilei revenite, în urna din care se extrag cele n bile aflîndu-se bile de culoare albă şi de culoare neagră, valorile variabilei aleatoare reprezentînd numărul de bile albe

din urna considerată, iar p fiind probabilitatea de a extrage o bilă albă din această urnă.

Am observat că atunci cînd experienţa se desfăşoară conform schemei bilei nerevenite, iar în urnă sînt t bile în total atunci variabila aleatoare asociată, Y, are media şi dispersia date de:

$$M(Y) = np$$
$$D(Y) = npq\frac{t-n}{t-1}$$

p şi q avînd aceeaşi semnificaţie, adică p este probabilitatea de a obţine o bilă albă din urnă, iar $q = 1 - p$ probabilitatea de a obţine o bilă neagră din aceeaşi urnă.

Să observăm că atunci cînd numărul total de bile din urnă, t ,este suficient de mare, în raport cu n, atunci $\frac{t-n}{t-1}$ este un număr apropiat de 1, iar dacă $t \to \infty$ atunci:

$$M(Y) = np$$
$$D(Y) \to npq$$

Cu alte cuvinte media şi dispersia variabilei aleatoare asociate unei experienţe conform schemei bilei revenite sau shemei bilei nerevenite pot fi considerate aceleaşi, atunci numărul elementelor colectivităţii din care se fac extracţiile este suficient de mare.

3. Repartiţia Poisson

Cele două clase de variabile aleatoare discrete considerate mai sus sînt formate din variabile aleatoare simple, adică cu un număr finit de valori. Vom considera acum o clasă de variabile aleatoare discrete, care au o mulţime numărabilă de valori.

Definiţia 7.7.3.

Fie (Ω, K, P) un cîmp de probabilitate. Se spune că o variabilă aleatoare de tip discret X, definită peste acest cîmp de probabilitate, are o repartiţie Poisson dacă tabloul său de repartiţie este de forma:

$$\begin{pmatrix} 0 & 1 & 2 & \dots & k & \dots \\ p_0 & p_1 & p_2 & \cdots & p_k & \cdots \end{pmatrix} \qquad 7.7.24$$

unde $p_k = \frac{\lambda^k}{k!} e^{-\lambda}$, $k = 0, 1, 2, \dots,$, λ fiind un număr real pozitiv.

O astfel de repartiţie este întîlnită în practică îndeosebi atunci cînd se urmăreşte studiul unor şiruri de aşteptare.

Ca şi în cazurile anterioare vom dovedi mai întîi că tabelul (7.7.24) este tabloul de repartiţie al unei variabile aleatoare, pentru orice $\lambda > 0$, adică

a) $\frac{\lambda^k}{k!} e^{-\lambda} \geq 0$, $k = 0, 1, 2, \dots$

b) $\sum_{k=0}^{\infty} \frac{\lambda^k}{k!} e^{-\lambda} = 1$

Deoarece $\lambda > 0$, $e^{-\lambda} > 0$, $k! > 0$, $k = 0, 1, 2, \dots$, prima condiţie este clar îndeplinită. Pentru a dovedi că şi cea de a doua condiţie este îndeplinită să observăm că $\sum_{k=0}^{\infty} \frac{\lambda^k}{k!}$ este dezvoltarea în serie Taylor a lui e^{λ}, adică:

$$e^{\lambda} = \sum_{k=0}^{\infty} \frac{\lambda^k}{k!}$$

Dar atunci putem scrie:

$$\sum_{k=0}^{\infty} \frac{\lambda^k}{k!} e^{-\lambda} = e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = e^0 = 1$$

şi deci tabloul (7.7.24) este tabloul de repartiţie alunei variabile aleatoare, pentru orice $\lambda > 0$.

Media unei variabile aleatoare cu repartiţia Poisson

Conform definiţiei mediei variabile aleatoare de tip discret, dacă X este o variabilă aleatoare cu repartiţie Poisson atunci:

$$M(X) = \sum_{k=0}^{\infty} k \frac{\lambda^k}{k!} e^{-\lambda} = e^{-\lambda} \sum_{k=1}^{\infty} \frac{\lambda^k}{(k-1)!} \qquad 7.7.25$$

Deoarece:

$$\sum_{k=1}^{\infty} \frac{\lambda^k}{(k-1)!} = \lambda \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} = \lambda \sum_{h=0}^{\infty} \frac{\lambda^h}{h!}$$

iar $\sum_{h=0}^{\infty} \frac{\lambda^h}{h!} = e^{\lambda}$, din (7.7.25) rezultă:

$$M(X) = e^{-\lambda} \lambda e^{\lambda} \qquad 7.7.26$$

sau echivalent:

$$M(X) = \lambda \qquad 7.7.27$$

Dispersia unei variabile aleatoare cu repartiţia Poisson

Deoarece $D(X) = M(X^2) - [M(X)]^2$ şi am arătat că pentru o variabilă aleatoare X, cu repartiţie Poisson, are loc $M(X) = \lambda$, pentru a calcula dispersia unei astfel de variabile aleatoare este suficient să mai determinăm $M(X^2)$.

Din cele arătate mai sus, putem scrie succesiv:

$$M(X^2) = \sum_{k=2}^{\infty} k^2 \frac{\lambda^k}{k!} e^{-\lambda} = e^{-\lambda} \sum_{k=1}^{\infty} k(k-1) \frac{\lambda^k}{k!} + e^{-\lambda} \sum_{k=1}^{\infty} k \frac{\lambda^k}{k!} = \qquad 7.7.28$$

$$= e^{-\lambda} \sum_{k=2}^{\infty} \frac{\lambda^k}{(k-2)!} + \lambda = \lambda^2 + \lambda$$

Prin urmare are loc:

$$D(X) = \lambda^2 + \lambda - \lambda^2 \qquad 7.7.29$$

sau echivalent:

$$D(X) = \lambda . \qquad 7.7.30$$

Repartiţia Poisson se mai numeşte şi *„legea evenimentelor rare"*, denumire care se justifică din următoarele considerente: Să presupunem că într-o urnă se află bile albe şi bile negre şi să notăm cu A evenimentul care se produce atunci cînd din urnă se extrage o bilă albă,iar cu p probabilitatea evenimentului A. Dacă din această urnă se extrag succesiv n bile, conform schemei bilei revenite şi asociem acestei experienţe o variabilă aleatoare care are ca valori numărul de bile albe extrase, această variabilă aleatoare are o repartiţie binomială, cu tabloul de repartiţie:

$$\begin{pmatrix} k \\ C_n^k p^k q^{n-k} \end{pmatrix}; \; k = 0, 1, \ldots, n$$

unde $q = 1 - p$. Deasemeni, am constatat că media acestei variabile aleatoare este np. Să presupunem acum că probabilitatea p, de extragere a unei bile albe din urnă, depinde de numărul n de bile extrase în cursul unei astfel de experienţe, astfel încît $\lim_{n\to\infty} np = \lambda$, unde λ este o constantă pozitivă. Dacă se întîmplă aşa ceva, atunci este clar că p se apropie de zero pe măsură ce n creşte, sau, altfel spus, probabilitatea evenimentului A tinde spre zero, adică putem spune că A este un „eveniment care se produce rar". Vom arăta că în aceste condiţii are loc relaţia:

$$\lim_{n\to\infty} C_n^k p^k q^{n-k} = \frac{\lambda^k}{k!} e^{-\lambda} \qquad 7.7.31$$

Într-adevăr,putem scrie succesiv:

$$C_n^k p^k q^{n-k} = \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k} = \frac{(np)^k}{k!} \frac{n!}{n^k (n-k)!} (1-p)^{n-k}$$

Dacă notăm $np = \lambda(n)$, (conform ipotezei avem $\lim_{n\to\infty} \lambda(n) = \lambda$), din relaţiile de mai sus urmează:

$$C_n^k p^k q^{n-k} = \frac{(np)^k}{k!} \frac{n(n-1)\ldots(n-k+1)}{n^k} \left[1 - \frac{\lambda(n)}{n}\right]^{n-k}$$

Dar $\lim_{n\to\infty} np = \lambda$, $\lim_{n\to\infty} \frac{n(n-1)\ldots(n-k+1)}{n^k} = 1$, iar:

$$\lim_{n\to\infty}\left[1-\frac{\lambda(n)}{n}\right]^{n-k}=\lim_{n\to\infty}\left[\left(1-\frac{\lambda(n)}{n}\right)^{-\frac{n}{\lambda(n)}}\right]^{-\lambda(n)\frac{n-k}{n}}=e^{-\lambda}$$

şi atunci:

$$\lim_{n\to\infty} C_n^k p^k q^{n-k}=\frac{\lambda^k}{k!}e^{-\lambda}$$

ceea ce dovedeşte relaţia (7.7.31).

§ 8. REPARTIŢII UZUALE DE TIP CONTINUU

Următoarele consideraţii sînt consacrate prezentării unor repartiţii de tip continuu, folosite în mod curent în statistica matematică.

1. Repartiţia normală standard

Definiţia 7.8.1.

Fie (Ω, K, P) un cîmp de probabilitate şi X o variabilă aleatoare de tip continuu definite peste acest cîmp de probabilitate. Vom spune că variabila aleatoare X are repartiţia standard, sau repartiţia normală $N(0,1)$, dacă funcţia sa de densitate, f, este dată prin:

$$f(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}};\ x\in\mathbb{R} \qquad 7.8.1$$

După cum se ştie, pentru ca funcţia f definită prin (7.8.1) să poată fi o funcţie de densitate pentru o variabilă aleatoare, aceasta trebuie să îndeplinească condiţiile:

$$\begin{aligned} &f(x)\geq 0\ ;\ (\forall)x\in\mathbb{R}\\ &\int_{-\infty}^{+\infty} f(x)\,dx=1 \end{aligned} \qquad 7.8.2$$

Prima dintre condiţii este evident îndeplinită deoarece $\frac{1}{\sqrt{2\pi}}>0$ iar

$e^{-\frac{x^2}{2}} > 0$, $(\forall) x \in \mathbb{R}$. Pentru a dovedi că şi cea de a doua condiţie din (7.8.2) este îndeplinită, vom ţine cont de faptul că integrala lui Gauss are valoarea $\frac{\sqrt{\pi}}{2}$, adică:

$$\int_0^{\infty} e^{-x^2} dx = \frac{\sqrt{\pi}}{2} \qquad 7.8.3$$

Ţinînd cont de proprietăţile integralei definite, putem scrie succesiv:

$$\int_{-\infty}^{+\infty} f(x)\,dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{x^2}{2}} dx = \frac{2}{\sqrt{2\pi}} \int_0^{\infty} e^{-\frac{x^2}{2}} dx \qquad 7.8.3'$$

iar după efectuarea substituţiei $x = t\sqrt{2}$ în ultima dintre aceste integrale, rezultă:

$$\int_{-\infty}^{+\infty} f(x)\,dx = \frac{2}{\sqrt{\pi}} \int_0^{\infty} e^{-x^2} dx \qquad 7.8.4$$

Din această relaţie şi (7.8.3) rezultă imediat:

$$\int_{-\infty}^{+\infty} f(x)\,dx = 1$$

Prin urmare funcţia f definită prin (7.8.1) este o funcţie de densitate de probabilitate.

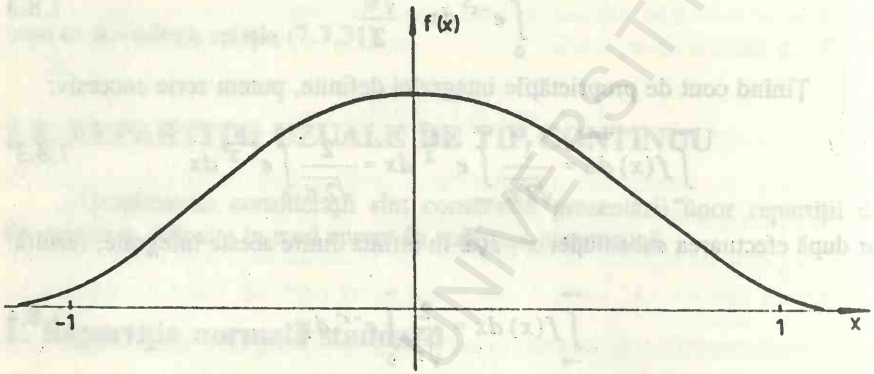
Forma graficului acestei funcţii, numit şi clopotul lui Gauss, este reprezentată în Fig.1 (pagina următoare).

Deoarece funcţia f definită prin (7.8.1) este o funcţie pară, $(f(-x) = f(x))$, rezultă că acest grafic este simetric faţă de axa ordonatelor, graficului avînd două puncte de inflexiune, de abscise ± 1.

Vom arăta că media şi dispersia pentru o variabilă aleatoare cu repartiţie normală standard sînt

$$M(X) = 0$$
$$D(X) = 1 \qquad 7.8.5$$

Într-adevăr, fie X o variabilă aleatoare cu repartiţie normală standard.



Conform definiţiei unei variabile aleatoare de tip continuu, urmează:

$$M(X) = \int_{-\infty}^{+\infty} x \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx \qquad 7.8.6$$

Deoarece funcţia de integrat în (7.8.6) este o funcţie impară, iar intervalul de integrare este simetric faţă de originea axelor rezultă că această integrală este egală cu zero.

$$\int_{-\infty}^{+\infty} x \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = 0$$

Deci $M(x) = 0$, ceea ce dovedeşte prima dintre relaţiile (7.8.5).

Din cauza faptului că pentru o variabilă aleatoare cu repartiţia normală standard are loc $M(X) = 0$ urmează că $D(X) = M(X^2)$. După cum se ştie, pentru o variabilă aleatoare, X, de tip continuu, cu funcţia de densitate f are loc:

BCU IASI - CENTRAL UNIVERSITY LIBRARY

$$M(X^2) = \int_{-\infty}^{+\infty} x^2 f(x)\,dx \qquad 7.8.7$$

iar în cazul cînd X are repartiţia normală standard:

$$M(X^2) = \int_{-\infty}^{+\infty} x^2 \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\,dx \qquad 7.8.8$$

Deoarece funcţia de integrat în (7.8.8) este o funcţie pară urmează:

$$M(X^2) = \frac{2}{\sqrt{2\pi}} \int_{0}^{+\infty} x^2 e^{-\frac{x^2}{2}}\,dx \qquad 7.8.9$$

şi aplicînd integralei din această relaţie formula de integrare prin părţi, una din părţi fiind $u = x$ iar cealaltă $dv = xe^{-\frac{x^2}{2}}\,dx$, obţinem:

$$\frac{2}{\sqrt{2\pi}} \int_{0}^{\infty} x^2 e^{-\frac{x^2}{2}}\,dx = \frac{2}{\sqrt{2\pi}} \int_{0}^{\infty} e^{-\frac{x^2}{2}}\,dx \qquad 7.8.10$$

Pe baza relaţiilor (7.8.3) şi a faptului că $\int_{-\infty}^{+\infty} f(x)\,dx = 1$ rezultă că $\frac{2}{\sqrt{2\pi}} \int_{0}^{\infty} e^{-\frac{x^2}{2}}\,dx = 1$ şi atunci, din (7.8.9) şi (7.8.10), urmează direct:

$$M(X^2) = 1. \qquad 7.8.11$$

Deci $D(X) = M(X^2) = 1$, ceea ce justifică şi a doua dintre relaţiile(7.8.5).

Funcţia lui Laplace

Pentru repartiţiile de tip continuu pe care le vom prezenta este necesară de multe ori cunoaşterea valorilor funcţiei de repartiţie corespunzătoare, pentru valori date ale argumentului. Funcţiile de densitate, prin intermediul cărora se dau aceste reparaţii în general, nu permit exprimarea funcţiei de repartiţie cu ajutorul funcţiilor elementare uzuale. Aceasta a determinat construirea unor tabele

de valori aproximative ale unor astfel de funcţii, sau ale unor funcţii prin intermediul cărora valorile funcţiilor de repartiţie corespunzătoare se determină foarte simplu. Pentru determinarea valorilor funcţiei de repartiţie a unei variabile aleatoare cu repartiţie normală standard, este utilizată funcţia integrală a lui Laplace.

Definiţia 7.8.2.

Vom numi funcţia lui Laplace funcţia $\Phi : \mathbb{R} \to \mathbb{R}$, definită prin:

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{t^2}{2}} dt \; ; \; (\forall) x \in \mathbb{R} \qquad 7.8.12$$

Dintre proprietăţile funcţiei lui Laplace:

P_1. Dacă x_1, x_2 atunci $\Phi(x_1) < \Phi(x_2)$, adică funcţia este o funcţie monoton crescătoare.

P_2. $\Phi(-x) = -\Phi(x)$; $(\forall) x \in \mathbb{R}$, adică funcţia Φ este o funcţie impară.

P_3. Dacă F este funcţia de repartiţie a unei variabile aleatoare cu repartiţie normală standard atunci:

$$F(x) = \frac{1}{2} + \Phi(x) \; ; \; (\forall) x \in \mathbb{R}$$

Demonstraţie:

P_1: Din definiţia funcţiei lui Laplace şi din faptul că $x_1 < x_2$ rezultă direct că are loc:

$$\Phi(x_2) = \Phi(x_1) + \frac{1}{\sqrt{2\pi}} \int_{x_1}^{x_2} e^{-\frac{t^2}{2}} dt \qquad 7.8.13$$

şi deoarece $\frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} > 0$, $(\forall) t \in \mathbb{R}$, iar $x_2 > x_1$ rezultă că $\frac{1}{\sqrt{2\pi}} \int_{x_1}^{x_2} e^{-\frac{t^2}{2}} dt > 0$.

Dar atunci, din (7.8.13), urmează:

$$\Phi(x_2) > \Phi(x_1) \qquad 7.8.14$$

ceea ce demonstrează proprietatea P_1.

BCU IASI / CENTRAL UNIVERSITY LIBRARY

P_2: Conform relaţiei de definiţie a funcţiei a lui Laplace, constatăm că are loc:

$$\Phi(-x) = \frac{1}{\sqrt{2\pi}} \int_0^{-x} e^{-\frac{t^2}{2}} dt \qquad 7.8.15$$

şi, dacă în integrala din (7.8.15) face substituţia $t = -u$, obţinem:

$$\Phi(-x) = -\frac{1}{\sqrt{2\pi}} \int_0^{x} e^{-\frac{t^2}{2}} dt \qquad 7.8.16$$

adică $\Phi(-x) = -\Phi(x)$ ceea ce justifică proprietatea P_2.

P_3. Pe baza definiţiei funcţiei de densitate pentru o variabilă aleatoare de tip continuu, rezultă că funcţia de repartiţie F, pentru o variabilă aleatoare cu repartiţie normală standard, este dată de:

$$F(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}} dt \; ; \; (\forall) x \in \mathbb{R} \qquad 7.8.17$$

Din această relaţie şi din proprietăţile integralei definite urmează direct:

$$F(x) = \int_{-\infty}^{0} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt + \int_0^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt \; ; \; (\forall) x \in \mathbb{R} \qquad 7.8.18$$

Deoarece $\int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt = 1$, iar funcţia de integrat este pară rezultă că $\int_{-\infty}^{0} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt = \frac{1}{2}$ şi atunci, din (7.8.18) şi definiţia funcţiei lui Laplace urmează nemijlocit:

$$F(x) = \frac{1}{2} + \Phi(x) \; ; \; (\forall) x \in \mathbb{R} \qquad 7.8.19$$

ceea ce dovedeşte proprietatea P_3.

Ţinînd cont de proprietatea P_3 a funcţiei lui Laplace se observă că valorile funcţiei de repartiţie pentru o variabilă aleatoare cu repartiţia normală stan-

dard se pot determina imediat ce se cunosc valorile corespunzătoare ale funcţiei lui Laplace. Din proprietatea P_3 a acestei funcţii se vede că pentru a cunoaşte valoarea lui Laplace pentru un x oarecare este suficient să se cunoască valorile acestei funcţii numai pentru valori pozitive ale argumentului. Există tabele de valori ale funcţiei lui Laplace, $\Phi(x)$, pentru valori pozitive ale argumentului x, (vezi tabelul 1). În cazul cînd este nevoie să se determine valoarea funcţiei $\Phi(x)$ pentru $x<0$ se ţine cont de faptul că $\Phi(-x) = -\Phi(x)$. De exemplu să determinăm, folosind tabelul 1, $\Phi(-2)$. Deoarece:

$$\Phi(-2) = -\Phi(2)$$

iar în tabelul 1 se găseşte $\Phi(2) \approx 0,47725$ rezultă că $\Phi(-2) \approx -0,47725$.

Vom da acum două exemple de utilizare a tabelelor de valori pentru funcţia lui Laplace în scopul obţinerii unor informaţii relativ la o variabilă aleatoare cu repartiţie normală standard.

Exemplul 7.8.1.

Fie X o variabilă aleatoare cu repartiţie normală $N(0,1)$. Să se determine probabilitatea ca variabila aleatoare X să ia valori în intervalul (1,3), adică $P(-1<X<3)$.

Pe baza proprietăţilor funcţiei de repartiţie a unei variabile aleatoare rezultă că are loc:

$$P(-1<X<3) = F(3) - F(-1)$$

şi deoarece $F(3) = \frac{1}{2} + \Phi(3)$, $F(-1) = \frac{1}{2} + \Phi(-1)$ rezultă:

$$P(-1<X<3) = \Phi(3) - \Phi(-1)$$

Din tabelul de valori ale funcţiei lui Laplace se găseşte $\Phi(3) \approx 0,49865$ şi $\Phi(1) \approx 0,34134$. Deoarece $\Phi(-1) = -\Phi(1)$ rezultă că

$$P(-1<X<3) \approx 0,49865 + 0,34134 \approx 0,83999$$

Exemplul 7.8.2.

Fie X o variabilă aleatoare cu repartiţie normală $N(0,1)$. Să determinăm

$z, z > 0$, astfel încît $P(|X| < z) = 0,8$. Deoarece $P(|X| < z) = P(-z < X < z)$, pe baza proprietăţilor funcţiei lui Laplace, obţinem:

$$P(|X| < z) = 2\Phi(z)$$

şi deci $P(|X| < z) = 0,8$ atunci cînd $2\Phi(z) = 0,8$, adică $\Phi(z) = 0,4$. Din tabelele de valori ale funcţiei lui Laplace se constată că $\Phi(z) = 0,39973$ pentru $z = 1,28$ şi $\Phi(z) = 0,40147$ pentru $z = 1,29$. Deoarece funcţia Φ este o funcţie crescătoare, prin interpolare se găseşte drept valoare aproximativă a soluţiei ecuaţiei $\Phi(z) = 0,4$ numărul $z_0 = 1,2815$, deci putem spune că $P(|X| < z) = 0,8$ pentru $z \approx 1,2815$.

2. Repartiţia normală generală

Definiţia 7.8.3.

Fie (Ω, K, P) un cîmp de probabilitate şi X o variabilă aleatoare de tip continuu definită peste acest cîmp de probabilitate. Se spune că variabila aleatoare X are o repartiţie normală generală, sau că are o repartiţie normală $N(a, \sigma)$, dacă funcţia sa de densitate, f, este dată prin:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-a)^2}{2\sigma^2}} \quad ; x \in \mathbb{R} \qquad 7.8.20$$

unde $a, \sigma \in \mathbb{R}$ şi $\sigma > 0$.

Să arătăm că funcţia $f: \mathbb{R} \to \mathbb{R}$, definită prin (7.8.20) este funcţia de densitate pentru o variabilă aleatoare.

După cum ştim aceasta înseamnă că funcţia f trebuie să îndeplinească condiţiile:

a) $f(x) \geq 0$, $(\forall) x \in \mathbb{R}$

b) $\int_{-\infty}^{+\infty} f(x)\,dx = 1$

Deoarece $\sigma > 0$, este evident că funcţia f definită prin (7.8.20) satisface condiţia a). Să dovedim că funcţia f îndeplineşte şi cea de a doua condiţie. Con-

form definiţiei funcţiei f, putem scrie:

$$\int_{-\infty}^{+\infty} f(x)\,dx = \int_{-\infty}^{+\infty} \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{(x-a)}{2\sigma^2}\,dx} \qquad 7.8.21$$

şi, după ce în integrala din membrul al doilea facem substituţia $\frac{x-a}{\sigma} = t$, obţinem:

$$\int_{-\infty}^{+\infty} f(x)\,dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{t^2}{2}}\,dt \qquad 7.8.22$$

şi deoarece am dovedit că funcţia, f, definită prin (7.8.20) este o funcţie de densitate rezultă că:

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{t^2}{2}}\,dt = 1$$

ceea ce arată că din (7.8.22) urmează direct:

$$\int_{-\infty}^{+\infty} f(x)\,dx = 1 \qquad 7.8.23$$

Prin urmare funcţia f, definită prin (7.8.20), este o funcţie de densitate de probabilitate.

Graficul funcţiei de densitate pentru o repartiţie normală $N(a, \sigma)$ este de forma reprezentată în figura 2. Se observă (şi se poate dovedi) că acest grafic este simetric faţă de perpendiculara în a pe axa absciselor, graficul fiind cu atît mai strîns în jurul acestei axe cu cît σ este mai mic. Cu linie întreruptă sînt desenate graficele funcţiei de densitate pentru o valoare mai mare şi o valoare mai mică pentru σ faţă de o valoare dată σ_0, valoarea lui a rămînînd neschimbată.

De asemeni se constată că tendinţa de apropiere a graficului de axa absciselor pe măsură ce x se depărtează de a, este foarte rapidă, (dacă $|x - a| < 3\sigma$ atunci $f(x) = < 0,003$).

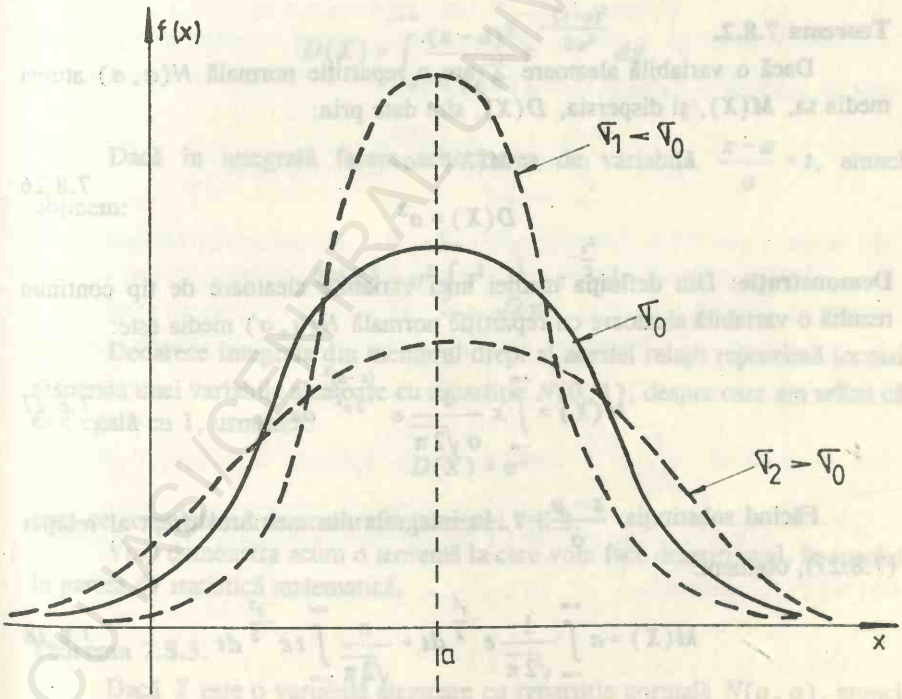
Teorema 7.8.1.

Funcţia de repartiţie, F, a unei variabile aleatoare cu repartiţie normală $N(a, \sigma)$ este de forma:

$$F(x) = \frac{1}{2} + \Phi\left(\frac{x-a}{\sigma}\right) ; (\forall) x \in \mathbb{R}$$

unde Φ este funcţia integrală a lui Laplace.

Demonstraţie: Deoarece o variabilă aleatoare cu repartiţia $N(a, \sigma)$ are funcţia de densitate de forma:



(7.8.20) urmează că funcţia sa de repartiţie F este dată prin:

$$F(x) = \int_{-\infty}^{x} \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{(t-a)^2}{2\sigma^2}} dt. \qquad 7.8.24$$

În urma schimbării de variabilă $\frac{t-a}{\sigma} = u$, în integrala din (7.8.24), urmează:

$$F(x) = \int_{-\infty}^{\frac{x-a}{\sigma}} \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du \; ; \; \forall x \in \mathbb{R} \qquad 7.8.25$$

şi deci $F(x) = \frac{1}{2} + \Phi\left(\frac{x-a}{\sigma}\right)$; $(\forall) x \in \mathbb{R}$, ceea ce demonstrează teorema.

Teorema 7.8.2.

Dacă o variabilă aleatoare X are o repartiţie normală $N(a, \sigma)$ atunci media sa, $M(X)$, şi dispersia, $D(X)$, sînt date prin:

$$\begin{aligned} M(X) &= a \\ D(X) &= \sigma^2 \end{aligned} \qquad 7.8.26$$

Demonstraţie: Din definiţia mediei unei variabile aleatoare de tip continuu rezultă o variabilă aleatoare cu repartiţie normală $N(a, \sigma)$ media este:

$$M(X) = \int_{-\infty}^{+\infty} x \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-a)^2}{2\sigma^2}} dx \qquad 7.8.27$$

Făcînd substituţia $\frac{x-a}{\sigma} = t$, în integrala din membrul drept al relaţiei (7.8.27), obţinem:

$$M(X) = a \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt + \frac{\sigma}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} t e^{-\frac{t^2}{2}} dt \qquad 7.8.28$$

Dar, cum am constatat la repartiţia normală $N(0, 1)$, prima integrală din

membrul al doilea al relaţiei (7.8.28) este egală cu 1, iar cea de a doua este egală cu zero aşa că, din (7.8.28), urmează:

$$M(X) = a$$

ceea ce demonstrează prima dintre relaţiile (7.8.26).

Pentru o variabilă aleatoare de tip continuu dispersia $D(X)$ este:

$$D(X) = \int_{-\infty}^{+\infty} [x - M(X)]^2 f(x)\, dx$$

unde $M(X)$ este media variabilei aleatoare,iar f este funcţia sa de densitate.În cazul unei variabile aleatoare X, cu repartiţie normală $N(a, \sigma)$, pentru care am dovedit că $M(X) = a$, dispersia va fi:

$$D(X) = \int_{-\infty}^{+\infty} \frac{(x-a)^2}{\sigma\sqrt{2\pi}} e^{-\frac{(x-a)^2}{2\sigma^2}}\, dx$$

Dacă în integrală facem schimbarea de variabilă $\frac{x-a}{\sigma} = t$, atunci obţinem:

$$D(X) = \sigma^2 \int_{-\infty}^{+\infty} t^2 \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}\, dt$$

Deoarece integrala din membrul drept al acestei relaţii reprezintă tocmai dispersia uneï variabile aleatoare cu repartiţie $N(0, 1)$, despre care am arătat că este egală cu 1, urmează:

$$D(X) = \sigma^2$$

ceea ce completează demonstraţia teoremei 7.8.2.

Vom demonstra acum o teoremă la care vom face deseori apel, în special în partea de statistică matematică.

Teorema 7.8.3.

Dacă X este o variabilă aleatoare cu repartiţie normală $N(a, \sigma)$, atunci variabila aleatoare normată asociată ei are repartiţia normală $N(0, 1)$.

Demonstraţie: Fie X o variabilă aleatoare cu repartiţia normală $N(a, \sigma)$. Să notăm cu F_1 funcţia de repartiţie a variabilei aleatoare X şi cu F_2 funcţia de repartiţie a variabilei aleatoare normate $Y = \dfrac{X - a}{\sigma}$.

Deoarece funcţia de densitate a variabilei aleatoare X este de forma (7.8.20) urmează că are loc:

$$F_1(x) = \int_{-\infty}^{x} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(t-a)^2}{2\sigma^2}} dt \; ; \; (\forall) x \in \mathbb{R}$$

şi, conform definiţiei funcţiei de repartiţie pentru o variabilă aleatoare, putem scrie:

$$F_2(x) = P\left(\frac{X - a}{\sigma} < x\right) = P(X < a + \sigma x)$$

ceea ce arată că $F_2(x) = F_1(a + \sigma x)$ şi atunci:

$$F_2(x) = \int_{-\infty}^{a+\sigma x} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(t-a)^2}{2\sigma^2}} dt \; ; \; (\forall) x \in \mathbb{R}$$

Dacă în integrală, facem substituţia $\dfrac{t - a}{\sigma} = u$ obţinem:

$$F_2(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt$$

ceea ce arată că $\dfrac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}$ este funcţie de densitate a variabilei aleatoare $Y = \dfrac{X - a}{\sigma}$ şi deci variabila aleatoare Y are repartiţie normală $N(0, 1)$. Teorema este astfel demonstrată.

BCU "ASICENTRA" UNIVERSITY LIBRARY

3. Repartiţia hi-pătrat

Definiţia 7.8.4.

Fie (Ω, K, P) un cîmp de probabiliatate şi X o variabilă aleatoare de tip continuu, definită peste acest cîmp de probabilitate. Vom spune că variabila aleatoare X are repartiţia hi-pătrat cu n grade de libertate dacă funcţia $f: \mathbf{R} \to \mathbf{R}$ definită prin:

$$f(x) = \begin{cases} 0 & ; x \le 0 \\ \dfrac{x^{\frac{n}{2}-1} e^{-\frac{x}{2}}}{2^{\frac{n}{2}} \Gamma\left(\frac{n}{2}\right)} & ; x > 0 \end{cases} \qquad 7.8.29$$

este funcţie de densitate pentru acestă variabilă aleatoare. În relaţia (7.8.29), n este un număr natural, iar Γ este funcţia gamma a lui Euler, adică:

$$\Gamma\left(\frac{n}{2}\right) = \int_0^{\infty} x^{\frac{n}{2}-1} e^{-x}\, dx$$

Vom dovedi că funcţia $f: \mathbf{R} \to \mathbf{R}$, definită prin (7.8.29) este o funcţie de densitate pentru o variabilă aleatoare.

Este clar că $f(x) \ge 0$, $(\forall) x \in \mathbf{R}$ şi atunci, pentru ca f să fie funcţie de densitate, rămîne să arătăm că este îndeplinită condiţia:

$$\int_{-\infty}^{+\infty} f(x)\, dx = 1$$

Din definiţia funcţiei f rezultă că putem scrie:

$$\int_{-\infty}^{+\infty} f(x)\, dx = \int_0^{\infty} \frac{1}{2^{\frac{n}{2}} \Gamma\left(\frac{n}{2}\right)} x^{\frac{n}{2}-1} e^{-\frac{x}{2}}\, dx$$

care, după substituţia $x = 2t$, devine:

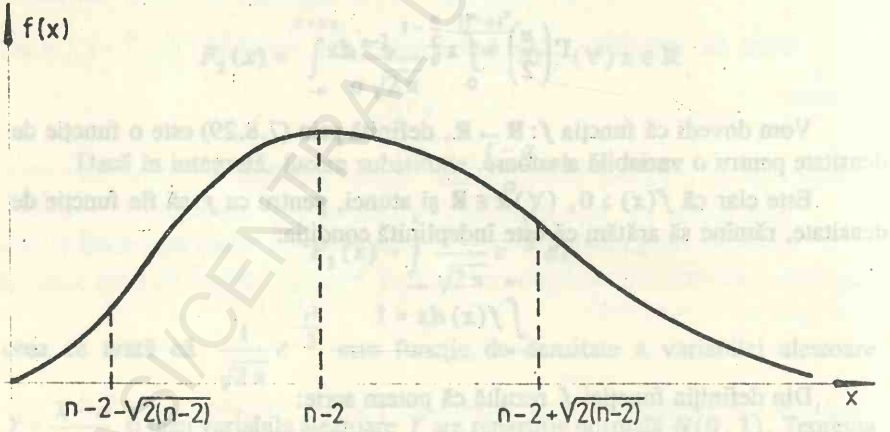
$$\int_{-\infty}^{+\infty} f(x) = \frac{1}{\Gamma\left(\frac{n}{2}\right)} \int_{0}^{\infty} t^{\frac{n}{2}-1} e^{-t}\, dt$$

de unde, ţinînd cont de definiţia funcţiei gamma, urmează:

$$\int_{-\infty}^{+\infty} f(x)\, dx = 1$$

ceea ce justifică afirmaţia că funcţia f, definită prin (7.8.29), este o funcţie de densitate de probabilitate.

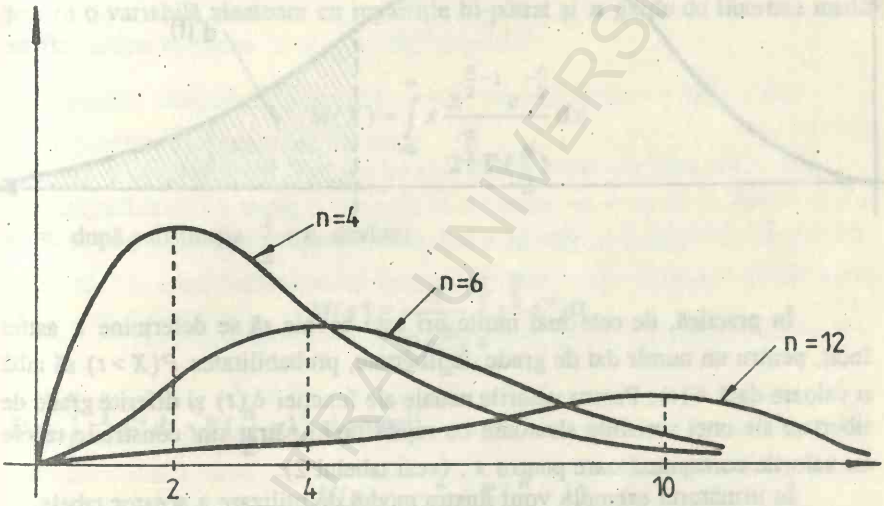
În cazul cînd $n \geq 3$, forma caracteristică a graficului funcţiei de densitate, pentru repartiţia hi-pătrat cu n grade de libertate, este reprezentată în figura 3.



În figura 4 de pe pagina următoare, este redat schematic modul cum se modifică graficul funcţiei de densitate odată cu modificarea numărului gradelor de libertate.

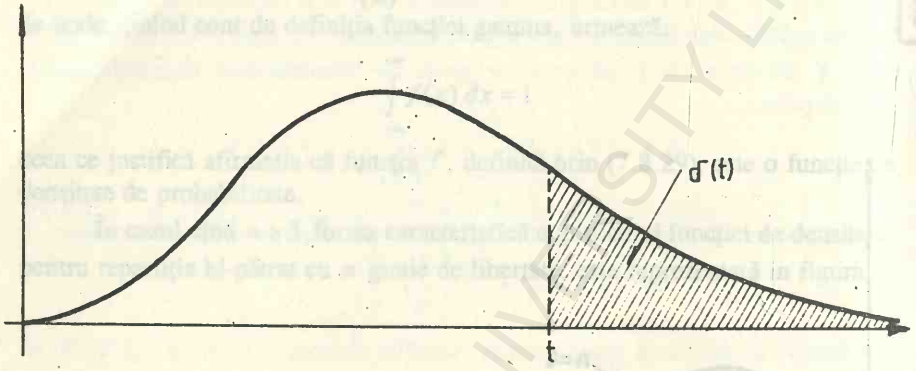
Funcţia de repartiţie a unei distribuţii hi-pătrat cu n grade de libertate va fi, conform definiţiei,

$$F(x) = P(X<x) = \int_0^x \frac{1}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)} t^{\frac{n}{2}-1} e^{-\frac{t}{2}} dt$$



În practică este adesea necesar să se cunoască valorile funcţiei $\delta(t) = P(X>t)$, unde X este o variabilă aleatoare cu repartiţie hi-pătrat şi un număr dat de grade de libertate. Pentru un număr de grade de libertate fixat, valoarea $\delta(t)$ reprezintă aria haşurată din figura 5 (vezi pagina următoare).

Din definiţia $\delta(t)$ este evident că $F(t) = 1 - \delta(t)$ ceea ce înseamnă că valorile funcţiei de repartiţie se pot determina foarte simplu dacă se cunosc valorile funcţiei $\delta(t) = P(X>t)$. Se observă că funcţia $\delta(t)$ este o funcţie descrescătoare în raport cu t, $\delta(0) = 1$ şi $\lim_{t\to\infty} \delta(t) = 0$.



În practică, de cele mai multe ori este nevoie să se determine t astfel încît, pentru un număr dat de grade de libertate, probabilitatea $P(X > t)$ să aibă o valoare dată $\delta(t)$. Pentru valorile uzuale ale funcţiei $\delta(t)$ şi diferite grade de libertate ale unei variabile aleatoare cu repartiţie hi-pătrat sînt construite tabele cu valorile corespunzătoare pentru t, (vezi tabelul 2).

În următorul exemplu vom ilustra modul de utilizare a acestor tabele.

Exemplul 7.8.3.

Fie X o variabilă aleatoare cu repartiţie hiperpătrat şi 5 grade de libertate. Se cere să se determine t_0 astfel încît probabilitatea ca variabila aleatoare să ia valori mai mari decît t_0, $P(X > t_0)$, să fie 0,05.

Deci trebuie să se determine t_0 astfel ca $\mathcal{P}(X > t_0) = 0{,}05$, unde X este o variabilă aleatoare cu repartiţie hi-pătrat şi 5 grade de libertate.

În tabelul 2, în coloana în care se găseşte probabilitatea $\delta = 0{,}05$ şi linia care dă numărul gradelor de libertate a variabilei aleatoare, $n = 5$, se află valoarea căutată a lui t_0, adică $t_0 = 11{,}07$.

Teorema 7.8.4.

Media şi dispersia pentru o variabilă X, cu repartiţia hi-pătrat şi n grade de libertate, sînt date de:

$$M(X) = n$$
$$D(X) = 2n \qquad 7.8.30$$

Demonstraţie: Conform definiţiei mediei unei variabile aleatoare de tip continuu, pentru o variabilă aleatoare cu repartiţie hi-pătrat şi n garde de libertate media va fi:

$$M(X) = \int_0^{\infty} x \frac{x^{\frac{n}{2}-1} e^{-\frac{x}{2}}}{2^{\frac{n}{2}} \Gamma(\frac{n}{2})} dx$$

care, după substituţia $\frac{x}{2} = t$, devine:

$$M(x) = \frac{2}{\Gamma(\frac{n}{2})} \int_0^{\infty} t^{\frac{n}{2}} e^{-t} dt$$

Dar $\int_0^{\infty} t^{\frac{n}{2}} e^{-t} dt = \Gamma(\frac{n}{2} + 1)$ şi atunci:

$$M(x) = \frac{2}{\Gamma(\frac{n}{2})} \Gamma(\frac{n}{2} + 1)$$

Din proprietăţile funcţiei gamma se ştie că are loc $\Gamma(a+1) = a\Gamma(a)$ şi deci $\Gamma(\frac{n}{2} + 1) = \frac{n}{2}\Gamma(\frac{n}{2})$ ceea ce înseamnă că $M(X) = n$. Deoarece $D(X) = M(X^2) - [M(X)]^2$, pentru a găsi dispersia unei variabile aleatoare, X, cu repartiţie hi-pătrat este suficient să mai determinăm $M(X^2)$.

$$M(X^2) = \int_0^{\infty} x^2 \frac{x^{\frac{n}{2}-1} e^{-\frac{x}{2}}}{2^{\frac{n}{2}} \Gamma(\frac{n}{2})} dx = \int_0^{\infty} \frac{x^{\frac{n}{2}+1} e^{-\frac{x}{2}}}{2^{\frac{n}{2}} \Gamma(\frac{n}{2})} dx$$

BCU CLUJ CENTRAL UNIVERSITY LIBRARY

După substituţia $\frac{x}{2} = t$ se constată că are loc:

$$M(X^2) = \frac{4}{\Gamma(\frac{n}{2})}\Gamma(\frac{n}{2}+2)$$

şi deoarece $\Gamma(\frac{n}{2}+2) = (\frac{n}{2}+1)\frac{n}{2}\Gamma(\frac{n}{2})$ urmează că:

$$M(X^2) = n(n+2)$$

Deci $D(X) = n(n+2) - n^2 = 2n$, ceea ce completează demonstraţia teoremei.

Se poate arăta că atunci cînd n este suficient de mare funcţia de repartiţie a unei variabile aleatoare cu repartiţia hi-pătrat şi n grade de libertate este bine aproximată de funcţia de repartiţie a unei variabile aleatore cu repartiţie normală $N(n, \sqrt{2n})$. Cu alte cuvinte, dacă X este o variabilă aleatoare cu repartiţie hi-pătrat şi n grade de libertate, atunci variabilela aleatoare normată asociată, $\frac{X-n}{\sqrt{2n}}$, se poate considera că urmează o repartiţie normală $N(0,1)$, atunci cînd n este suficient de mare. Practic se constată că aproximarea este suficient de bună pentru $n > 30$, motiv pentru care tabelele de valori ale funcţiei $P(X > t)$, unde X este o variabilă aleatoare cu repartiţie hi-pătrat şi n grade de libertate, se face numai pentru $n \le 30$.

Dacă $n > 30$ atunci valoarea probabilităţii $P(X > t)$ se aproximează cu $\frac{1}{2} - \Phi\left(\frac{t-n}{\sqrt{2n}}\right)$, unde Φ este funcţia lui Laplace. De exemplu, dacă X este o variabilă aleatoare cu repartiţie hi-pătrat şi 50 grade de libertate şi vrem să găsim t_0 astfel încît $P(X > t_0) = 0{,}05$, atunci această ecuaţie cu ecuaţia $\frac{1}{2} - \Phi\left(\frac{t_0 - 50}{\sqrt{2.50}}\right) = 0{,}05$.

Din tabelele de valori ale funcţiei Laplace, constatăm că $\Phi(z) = 0{,}45$ pentru $z \approx 1{,}645$. Prin urmare $P(X > t_0) = 0{,}05$ dacă t_0 verifică ecuaţia $\frac{t_0 - 50}{\sqrt{100}} = 1{,}645$,

de unde obţinem $t_0 = 66,45$.

Un alt rezultat util în practică, pe care aici îl vom enunţa numai, este exprimat în următoarea teoremă:

Teorema 7.8.5.

Dacă variabilele aleatoare $X_i, i = 1,2,\ldots,n$, sînt independente şi urmează o repartiţie normală $N(0,1)$, atunci variabila aleatoare $X = \sum_{i=1}^{n} X_i^2$ are repartiţie hi-pătrat cu n grade de libertate.

4. Repartiţia t (Repartiţia Student)

Definiţia 7.8.5.

Se spune că o variabilă aleatoare de tip continuu, X, urmează o repartiţie t, sau repartiţie Student, cu n grade de libertate, dacă funcţia $f: \mathbb{R} \to \mathbb{R}$ definită prin:

$$f(x) = \frac{1}{\sqrt{n}\beta\left(\frac{1}{n}, \frac{n}{2}\right)} \left(1 + \frac{x^2}{n}\right)^{-\frac{n+1}{2}}; x \in \mathbb{R} \qquad 7.8.31$$

este funcţia de densitate pentru variabila aleatoare X.

În relaţia de definiţie a funcţiei f prin $\beta(\frac{1}{2}, \frac{n}{2})$ se înţelege fucţia beta a lui Euler, $\beta(a,b)$, pentru $a = \frac{1}{2}$, $b = \frac{n}{2}$, adică:

$$\beta(\frac{1}{2}, \frac{n}{2}) = \int_0^1 x^{\frac{1}{2}-1} (1 - x^{\frac{n}{2}-1}) dx$$

Din relaţia (7.8.31) este clar că $f(x) > 0, (\forall) x \in \mathbb{R}$, şi atunci pentru a dovedi că funcţia f, definită prin această relaţie, este o funcţie de densitate de probabilitate trebuie să mai arătăm că are loc:

$$\int_{-\infty}^{+\infty} f(x)\,dx = 1$$

Deoarece funcţia f definită prin (7.8.31) este o funcţie pară, $f(-x)) = f(x)$, $(\forall)\, x \in \mathbb{R}$, rezultă:

$$\int_{-\infty}^{+\infty} f(x)\,dx = 2\int_{0}^{\infty} \frac{1}{\sqrt{n}\,\beta\left(\frac{1}{2}, \frac{n}{2}\right)} \left(1 + \frac{x^2}{n}\right)^{-\frac{n+1}{2}} dx$$

În urma schimbării de variabilă $y = \dfrac{x^2}{n + x^2}$, în integrala din membrul al doilea al relaţiei precedente, obţinem:

$$\int_{-\infty}^{+\infty} f(x)\,dx = 2\,\frac{1}{\sqrt{n}\,\beta\left(\frac{1}{2}, \frac{n}{2}\right)} \int_{0}^{1} y^{-\frac{1}{2}} (1-y)^{\frac{n}{2}-1} \frac{\sqrt{n}}{2}\,dy$$

şi ţinînd cont de definiţia funcţiei beta rezultă imediat:

$$\int_{-\infty}^{+\infty} f(x)\,dx = 1$$

ceea ce dovedeşte că funcţia f definită prin (7.8.31) este o funcţie de densitate de probabilitate.

Forma graficului funcţiei de densitate pentru o repartiţie t cu n grade de libertate este redată în figura 6 de pe pagina următoare (cu linie întreruptă este reprezentat graficul de densitate pentru repartiţia normală standard).

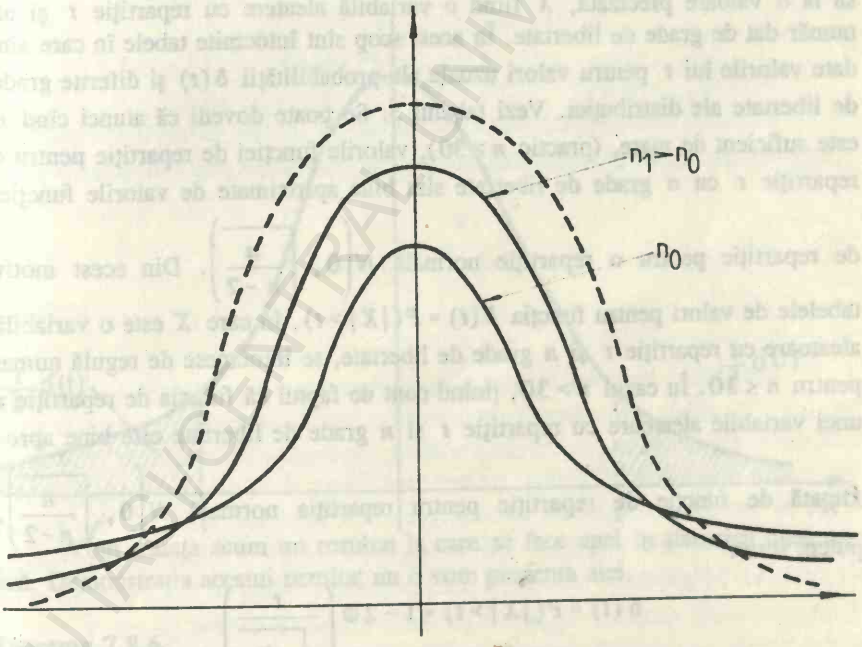
Se observă că pe măsură ce n creşte graficul funcţiei de densitate pentru repartiţia t se apropie din ce în ce mai mult de graficul funcţiei de densitate pentru repartiţia normală standard.

Deoarece funcţia de densitate pentru o variabilă aleatoare, X, cu repartiţie t şi n grade de libertate este o funcţie pară, rezultă imediat că media unei astfel de variabile aleatoare este $M(X) = 0$. Pentru a determina dispersia unei variabile aleatoare, X, cu repartiţia t şi n grade de libertate vom deter-

mina $M(X^2)$, care conform definiţiei funcţiei de densitate pentru o variabilă aleatoare cu repartiţia t şi n grade de libertate este:

$$M(X^2) = \int_{-\infty}^{+\infty} x^2 \frac{1}{\sqrt{n}\,\beta\left(\frac{1}{2}, \frac{n}{2}\right)} \left(1 + \frac{x^2}{n}\right)^{-\frac{n+1}{2}} dx =$$

$$= 2\int_{0}^{\infty} x^2 \frac{1}{\sqrt{n}\,\beta\left(\frac{1}{2}, \frac{n}{2}\right)} \left(1 + \frac{x^2}{n}\right)^{-\frac{n+1}{2}} dx$$



În urma substituţiei $y = \frac{x^2}{n + x^2}$ în această din urmă integrală şi, ţinînd cont de proprietăţile funcţiei β, constatăm că:

BCU "LUCIAN BLAGA" UNIVERSITY LIBRARY

$$M(X^2) = \frac{n}{n-2}$$

Deoarece $M(X) = 0$, rezultă că pentru o variabilă aleatoare, X, cu repartiţie t şi n grade de libertate dispersia este dată de:

$$D(X) = \frac{n}{n-2}$$

Relativ la repartiţia t, în problemele practice este deseori necesară determinarea unui număr real t astfel încît funcţia δ definită prin:

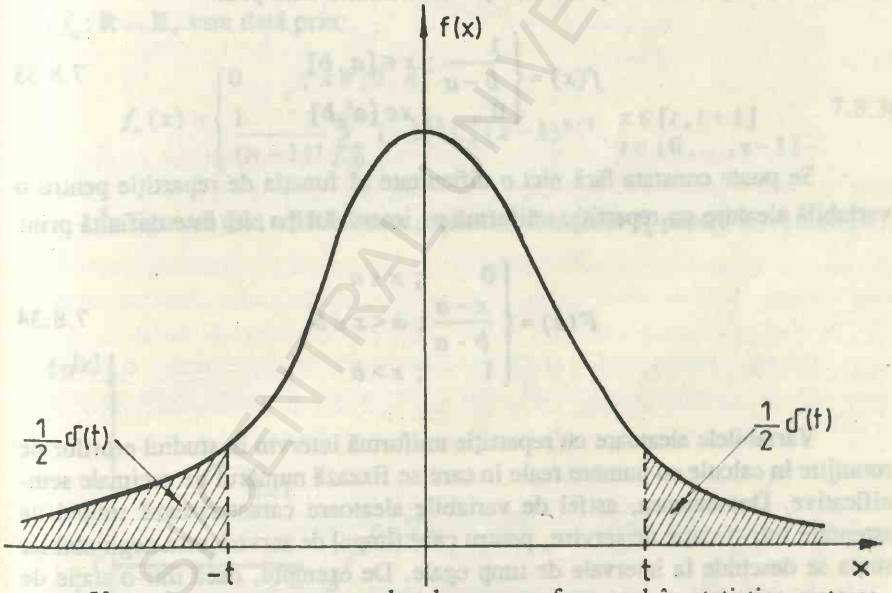
$$\delta(t) = P(|X| > t), \; t \geq 0 \qquad 7.8.32$$

să ia o valoare precizată, X fiind o variabilă aleatere cu repartiţie t şi un număr dat de grade de libertate. În acest scop sînt întocmite tabele în care sînt date valorile lui t pentru valori uzuale ale probabilităţii $\delta(t)$ şi diferite grade de libertate ale distribuţiei. Vezi tabelul 3. Se poate dovedi că atunci cînd n este suficient de mare, (practic $n > 30$), valorile funcţiei de repartiţie pentru o repartiţie t cu n grade de libertate sînt bine aproximate de valorile funcţiei de repartiţie pentru o repartiţie normală $N\left(0, \sqrt{\frac{n}{n-2}}\right)$. Din acest motiv tabelele de valori pentru funcţia $\delta(t) = P(|X| > t)$, în care X este o variabilă aleatoare cu repartiţie t şi n grade de libertate, se întocmesc de regulă numai pentru $n \leq 30$. În cazul $n > 30$, ţinînd cont de faptul că funcţia de repartiţie a unei variabile aleatoare cu repartiţie t şi n grade de libertate este bine aproximată de funcţie de repartiţie pentru repartiţia normală $N\left(0, \sqrt{\frac{n}{n-2}}\right)$, putem scrie:

$$\delta(t) = P(|X| > t) \approx 1 - 2\Phi\left(\frac{t}{\sqrt{\frac{n}{n-2}}}\right)$$

Observăm că pentru t fixat, $t \geq 0$, valoarea funcţiei $\delta(t)$ reprezintă aria haşurată din figura 7, graficul fiind al funcţiei de densitate pentru o repartiţie t cu n grade de libertate. Datorită faptului că funcţia de densitate pentru o repartiţie t cu n grade de libertate este o funcţie pară, putem constata imediat că legătura dintre funcţia de repartiţie F, a unei astfel de variabile aleatoare şi funcţia δ definită prin (7.8.32) este dată prin:

$$P(X<x) = F(x) = \begin{cases} \dfrac{\delta(-x)}{2}\ ; & x \leq 0 \\ 1 - \dfrac{\delta(x)}{2}\ ; & x > 0 \end{cases}$$



Vom enunţa acum un rezultat la care se face apel în statistica matematică. Demonstraţia acestui rezultat nu o vom prezenta aici.

Teorema 7.8.6.

Dacă variabilele aleatoare X şi Y sînt astfel încît X are repartiţia normală $N(0,1)$, iar Y are o repartiţie hi-pătrat cu n grade de libertate,

atunci variabila aleatoare $Z = \dfrac{X}{\sqrt{\dfrac{Y}{n}}}$ are o repartiţie t cu n grade de libertate.

5. Repartiţia uniformă

Definiţia 7.8.6.

Se spune că o variabilă aleatoare X are repartiţie uniformă pe un interval $[a, b]$ dacă este de tip continuu şi are funcţia de densitate dată prin:

$$f(x) = \begin{cases} \dfrac{1}{b-a}; & x \in [a,b] \\ 0 \quad\quad ; & x \notin [a,b] \end{cases} \qquad 7.8.33$$

Se poate constata fără nici o dificultate că funcţia de repartiţie pentru o variabilă aleatore cu repartiţie uniformă pe intervalul $[a, b]$ este definită prin:

$$F(x) = \begin{cases} 0 \quad\quad ; & x \le a \\ \dfrac{x-a}{b-a}; & a < x \le b \\ 1 \quad\quad ; & x > b \end{cases} \qquad 7.8.34$$

Variabilele aleatoare cu repartiţie uniformă intervin în studiul erorilor de rotunjire în calcule cu numere reale în care se fixează numărul de zecimale semnificative. Deasemenea, astfel de variabile aleatoare caracterizează timpul de aşteptare într-o staţie de servire, pentru care timpul de servire este neglijabil iar staţia se deschide la intervale de timp egale. De exemplu, dacă într-o staţie de metrou ramele sosesc cu regularitate la intervale de cîte t minute atunci timpul de aşteptare (exprimat în minute) al unui potenţial pasager se poate considera ca fiind o variabilă aleatoare cu repartiţie uniformă pe intervalul $[0, t]$.

Media şi dispersia unei variabile aleatoare cu repartiţia uniformă pe intervalul $[a, b]$ se poate determina cu uşurinţă. Prin calculul direct obţinem:

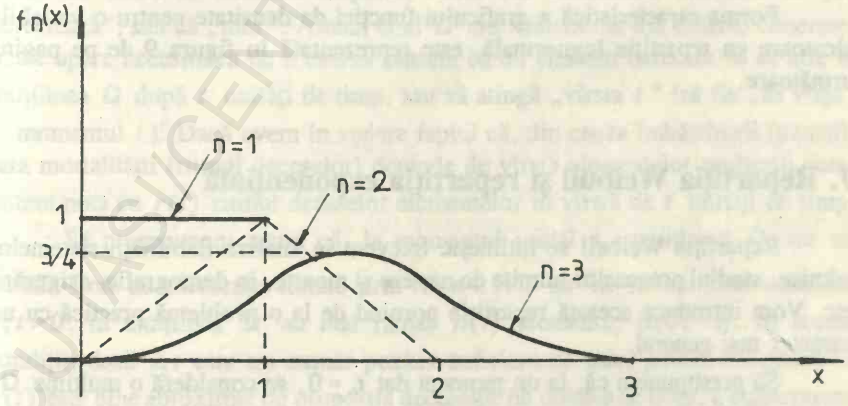
$$M(X) = \frac{a+b}{2}$$

$$D(X) = \frac{(b-a)^2}{12} \qquad 7.8.35$$

Menţionăm următoarea proprietate importantă a variabilelor aleatoare cu repartiţie uniformă. Dacă X_i, $i = 1, 2, \ldots, n$, sînt variabile aleatoare independente, cu repartiţia uniformă pe intervalul $[0, 1]$ atunci suma lor $Y = X_1 + \ldots + X_n$ are o repartiţie care, pe măsură ce n creşte, se apropie forte repede de repartiţia unei variabile normale. Mai precis, funcţia de densitate a variabilei aleatoare Y, $f_n : \mathbb{R} \to \mathbb{R}$, este dată prin:

$$f_n(x) = \begin{cases} 0 & ; x \notin [0, n] \\ \dfrac{1}{(n-1)!} \displaystyle\sum_{k=0}^{y} (-1)^k C_n^k (x-k)^{n-1} & , \begin{matrix} x \in [t, t+1] \\ t \in \{0, \ldots, n-1\} \end{matrix} \end{cases} \qquad 7.8.36$$

În figura 8 este redată forma graficelor funcţiilor f_n pentru $n \in \{1, 2, 3\}$.



6. Repartiţia lognormală

Definiţia 7.8.7.

Spunem că o variabilă aleatoare X are o repartiţie lognormală dacă variabila aleatoare $Y = \ln X$ are o repartiţie normală (în particular aceasta înseamnă că variabila aleatoare X are numai valori pozitive). Dacă presupunem că variabila aleatoare $Y = \ln X$ are o repartiţie normală $N(\ln a, \sigma^2)$, $a > 0, \sigma > 0$, atunci, procedînd la fel ca în demonstraţia teoremei 7.5.6, se poate constata că funcţia de densitate a variabilei aleatoare X este funcţia $f: \mathbb{R} \to \mathbb{R}$, definită prin:

$$f(x) = \begin{cases} 0 & , x \le 0 \\ \dfrac{1}{\sigma x \sqrt{2\pi}} e^{-\frac{(\ln x - \ln a)^2}{2\sigma^2}} & , x > 0 \end{cases} \qquad 7.8.37$$

Prin calcul direct, se poate determina cu uşurinţă media şi dispersia unei variabile aleatoare cu repartiţia lognormală şi obţinem:

$$M(X) = a e^{\frac{\sigma^2}{2}} \qquad 7.8.38$$

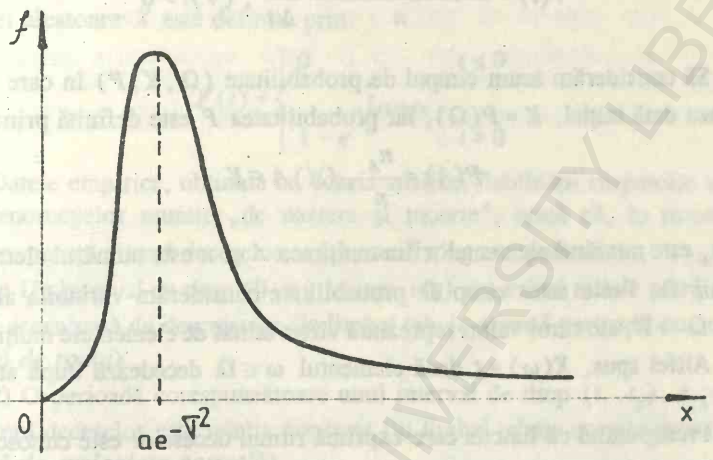
$$D(X) = a^2 e^{\sigma^2} (e^{\sigma^2} - 1)$$

Forma caracteristică a graficului funcţiei de densitate pentru o variabilă aleatoare cu repartiţie lognormală, este reprezentată în figura 9 de pe pagina următoare.

7. Repartiţia Weibull şi repartiţia exponenţială

Repartiţia Weibull se întîlneşte frecvent în analiza fiabilităţii sistemelor tehnice, studiul proceselor numite de naştere şi moarte, în demografie, asigurării etc. Vom introduce această repartiţie pornind de la o problemă practică cu un caracter mai general.

Să presupunem că, la un moment dat $t_0 = 0$, se consideră o mulţime Ω, constituită din elemente concrete (indivizi) care se disting prin faptul că fiecare



din ele au o serie de proprietăţi specifice. După trecerea a t unităţi de timp de la momentul t_0, unele din momentele aflate iniţial în mulţimea Ω îşi pierd o parte (eventual toate) din proprietăţile pe baza cărora au fost grupate în mulţimea Ω şi, drept urmare, sînt eliminate din această mulţime. În astfel de situaţii vom spune că odată cu trecerea timpului, unele din elementele mulţimii Ω „se defectează", sau că „mor". Atunci cînd Ω este constituită din obiecte concrete, poate apare necesitatea de a evalua şansele ca un element oarecare să se afle în mulţimea Ω după t unităţi de timp, sau să atingă „vîrsta t" (să fie „în viaţă" la momentul t). Dacă avem în vedere faptul că, din cauza îmbătrînirii (uzurii), rata mortalităţii (ritmul deceselor) depinde de vîrsta elementelor mulţimii date, putem nota cu $r(t)$ ritmul deceselor elementelor în vîrstă de t unităţi de timp.

Să presupunem acum că, la momentul iniţial t_0, mulţimea Ω are un număr finit de elemente notate prin $n, n \in \mathbb{N}^*$, iar la un moment oarecare $t, t > 0$, în mulţimea Ω au mai rămas $n(t)$ elemente, $n(0) = n$. În aceste condiţii, dacă Δt este un număr pozitiv suficient de mic, putem considera că $r(t)$ este bine aproximat de proporţia deceselor pe unitate de timp, a elementelor mulţimii care ating vîrsta t, adică:

$$r(t) \simeq \frac{n(t) - n(t+\Delta t)}{n} \frac{1}{\Delta t}, \quad (\forall) t > 0 \qquad 7.8.39$$

Să considerăm acum cîmpul de probabilitate (Ω, K, P) în care Ω este mulţimea dată iniţial, $K = P(\Omega)$, iar probabilitatea P este definită prin:

$$P(A) = \frac{n_A}{n}, \quad (\forall) A \in K \qquad 7.8.40$$

unde n_A este numărul elementelor din mulţimea A şi n este numărul elementelor mulţimii Ω. Peste acest cîmp de probabilitate considerăm variabila aleatoare X, $X: \Omega \to \mathbb{R}$, ale căror valori reprezintă vîrsta atinsă de elementele mulţimii Ω.

Altfel spus, $X(\omega) = t$ dacă elementul $\omega \in \Omega$ decedează după atingerea vîrstei t.

Presupunînd că funcţia care exprimă ritmul deceselor este cunoscută, ne propunem să determinăm funcţia de repartiţie a variabilei aleatoare X.

Dacă notăm cu F funcţia de repartiţie a acestei variabile aleatoare, atunci este clar că $P(X < t) = 0$ pentru orice $t \le 0$, iar pentru $t > 0$, din (7.8.40) rezultă:

$$F(t) = \mathcal{P}(X < t) = \frac{n - n(t)}{n}, \quad (\forall) t > 0$$

Pe de altă parte, din această ultimă relaţie urmează:

$$r(t) \simeq \frac{F(t+\Delta t) - F(t)}{\Delta t} \cdot \frac{1}{1 - F(t)}$$

pentru oricare $t > 0$ şi oricare Δt pozitiv, suficient de mic.

În ipoteza că eroarea comisă prin aproximarea tinde la zero odată cu Δt, iar funcţia F este derivabilă pe intervalul $(0, +\infty)$, trecînd la limită rezultă:

$$\frac{F'(t)}{1 - F(t)} = r(t), \quad (\forall) t > 0$$

Deoarece $F(0) = 0$, în urma rezolvării ecuaţiei diferenţiale obţinem:

$$F(t) = 1 - e^{-\int_0^t r(t)\,du}, \quad (\forall) t > 0$$

În concluzie, în ipotezele specificate mai înainte, funcţia de repartiţie a variabilei aleatoare X este definită prin:

$$F(t) = \begin{cases} 0 & ; t \le 0 \\ 1 - e^{-\int_0^t r(t)du} & ; t > 0 \end{cases}$$

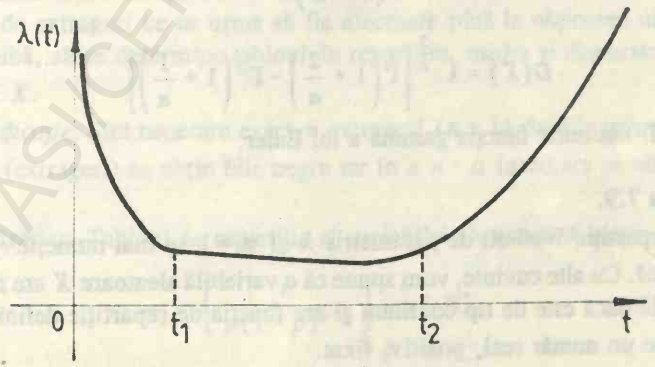
Datele empirice, obţinute cu ocazia analizei fiabilităţii sistemelor tehnice sau a fenomenelor numite „de naştere şi moarte", arată că, în majoritatea cazurilor, pentru ritmul deceselor, $r(t)$, se disting trei perioade:

1) Un interval de timp $(0, t_1)$ în care $r(t)$ are valori relativ mari, dar cu tendinţă accentuată de descreştere (în limbaj tehnic această perioadă este numită *perioadă de rodaj*).

2) O perioadă corespunzătoare unui interval de timp (t_1, t_2), $t_2 > t_1$, în care ritmul deceselor este relativ constant (în limbaj tehnic aceasta este numită *perioadă de exploatare normală*).

3) O perioadă corespunzătoare unui interval de timp $(t_2, +\infty)$, în care $r(t)$ se caracterizează printr-o tendinţă accentuată de creştere (această perioadă este numită *perioadă de uzură şi îmbătrînire*). Forma caracteristică a graficului funcţiei $r(t)$, corespunzătoare celor trei perioade menţionate mai înainte, este evidenţiată în figura 10. Fie λ şi α două numere reale fixate ($\lambda > 0$, $\alpha > 0$) şi funcţia $r: (0, +\infty) \to \mathbb{R}$, definită prin:

$$r(t) = \lambda\, \alpha\, t^{\alpha - 1}, \ (\forall)\, t > 0 \qquad 7.8.43$$



Dacă se consideră reprezentarea grafică a funcţiei r definită prin (7.8.43), atunci se poate observa că pentru $\alpha < 1$, $\alpha = 1$ şi respectiv $\alpha > 1$ graficul funcţiei r ar corespunde celor trei perioade reprezentative pentru ritmul deceselor.

Vom da acum definiţia unei variabile aleatoare cu repartiţie Weibull.

Definiţia 7.8.8.

Spunem că o variabilă aleatoare X are o repartiţie Weibull, de parametrii λ şi α, dacă variabila aleatoare X este de tip continuu şi are funcţia de repartiţie $F: \mathbb{R} \to \mathbb{R}$ definită prin:

$$F(x) = \begin{cases} 0 & ; x \le 0 \\ 1 - e^{-\lambda x^{\alpha}} & , x > 0 \end{cases} \qquad 7.8.44$$

Se observă că funcţia de repartiţie a unei variabile aleatoare cu repartiţie Weibull se obţine din precedenta.

Deoarece funcţuia de repartiţie a unei variabile aleatoare cu repartiţie Weibull este derivabilă, rezultă că funcţia de densitate este de forma:

$$f(x) = \begin{cases} 0 & ; x \le 0 \\ \lambda \alpha x^{\alpha - 1} e^{-\lambda x^{\alpha}} & ; x > 0 \end{cases}$$

Media şi dispersia unei variabile aleatoare cu repartiţie Weibull se pot determina fără dificultate, obţinîndu-se:

$$M(X) = \lambda^{-\frac{1}{\alpha}} \Gamma\left(1 + \frac{1}{\alpha}\right)$$

$$D(X) = \lambda^{-\frac{2}{\alpha}} \left[\Gamma\left(1 + \frac{2}{\alpha}\right) - \Gamma^2\left(1 + \frac{1}{\alpha}\right)\right]$$

unde prin Γ am notat funcţia gamma a lui Euler.

Observaţia 7.9.

O repartiţie Weibull de parametrii λ şi $\alpha = 1$ se mai numeşte *repartiţie exponenţială*. Cu alte cuvinte, vom spune că o variabilă aleatoare X are repartiţie exponenţială dacă este de tip continuu şi are funcţia de repartiţie definită prin: unde λ este un număr real, pozitiv, fixat.

$$F(x) = \begin{cases} 0 & ; x \le 0 \\ 1 - e^{-\lambda x} & ; x > 0 \end{cases}$$

Dacă funcţia de repartiţie a unei variante aleatoare cu repartiţie exponenţială este definită ca mai sus, rezultă că funcţia sa de densitate este de forma:

$$f(x) = \begin{cases} 0 & ; x \le 0 \\ \lambda e^{-\lambda x} & ; x > 0 \end{cases}$$

Prin calcul direct rezultă că media şi dispersia unei variabile aleatoare cu repartiţie exponenţială sînt date prin:

$$M(X) = \frac{1}{\lambda}, \; D(X) = \frac{1}{\lambda^2}$$

Notăm faptul că repartiţia exponenţială se întîlneşte, printre altele, în problemele de aşteptare. Dacă într-o staţie se servire ritmul servirilor este constant, atunci timpul de aşteptare al unui solicitant sosit la întîmplare în staţia de servire urmează o distribuţie exponenţială de parametru egal cu ritmul servirilor.

Exerciţii:

1. (*Repartiţie geometrică*). Într-o urnă se află bile albe şi bile negre, proporţia de bile albe existente în urnă fiind egală cu p. Din această urnă se extrage la întîmplare, cu revenire, cîte o bilă pînă cînd se obţine o bilă de culoare albă. Dacă notăm cu X variabila aleatoare ale cărei valori reprezintă numărul de extrageri ce ar urma să fie efectuate pînă la obţinerea unei bile de culoare albă, să se determine tabloul de repartiţie, media şi dispersia variabilei aleatoare X.

Indicaţie: Sînt necesare exact n extrageri ($n \ge 1$) dacă în primele $(n-1)$ încercări (extrageri) se obţin bile negre iar în a $n-a$ încercare se obţine o bilă albă.

Răspuns: Tabloul de repartiţie al variabilei aleatoare X este:

$$\begin{pmatrix} n \\ p(1-p)^{n-1} \end{pmatrix}, \; n \in \mathbb{N}^*$$

$$M(X) = \frac{1}{p}\ ,\ D(X) = \frac{(1-p)}{p_2}$$

2. (*Repartiţia binomială negativă*). Fie o experienţă (probă) în care un eveniment dat A se produce cu probabilitatea $\mathcal{P}(A) = p$. Se repetă experienţa (se fac probe) pînă cînd se constată că evenimentul A s-a produs de k ori. Dacă probabilitatea producerii evenimentului A în fiecare probă este egală cu p (rezultatul obţinut în fiecare probă este independent de rezultatele probelor anterioare), să se determine tabloul de repartiţie, media şi dispersia variabile aleatorie X, ale cărei valori reprezintă numărul de probe necesare pentru ca evenimentul A să se producă exact de k ori (k este fixat).

Indicaţie. Sînt necesare n probe dacă în primele $n-1$ probe evenimentul A se produce de $(k-1)$ ori, iar în a $n-a$ probă se produce evenimentul A. Pentru calculul mediei şi dispersiei se va ţine cont de egalitatea

$$\frac{1}{(1-x)^t} = \sum_{n=0}^{\infty} x^n C_{n+t-1}^{t-1}\ ;\ (\forall)x \in (-1,1)\ ,\ (\forall)t \in \mathbb{N}^*$$

Răspuns: Tabloul de repartiţie al variabilei aleatoare X este

$$\begin{pmatrix} n \\ p_n \end{pmatrix};\ n \in \mathbb{N}^*\ ,\ n \geq k$$

unde $p_n = C_{n-1}^{k-1} p^k (1-p)^{n-k}$, $(\forall n \in \mathbb{N},\ n \geq k)$.

$$M(X) = \frac{k}{p}\ ,\ D(X) = \frac{k(1-p)}{p^2}$$

3. Dacă X este o variabilă aleatoare cu repartiţie uniformă pe intervalul $[a, b]$, să se arate că variabila aleatoare $Y = \dfrac{(X-a)}{(b-a)}$ are repartiţie uniformă pe intervalul $[0, 1]$.

4. (*Repartiţia Pareto*). Se spune că o variabilă aleatoare X are o repartiţie Pareto dacă funcţia de repartiţie a acesteia este de forma:

$$F(x) = \begin{cases} 1 - \left(\dfrac{c}{x}\right)^{\alpha} & ;\ x > c \\ 0 & ;\ x \leq c \end{cases}$$

unde c şi α sînt numere reale fixate $c>0$, $\alpha>0$. Să se determine funcţia de densitate, media şi dispersia unei variabile aleatoare cu repartiţie Pareto.

Răspuns:

Dacă $\alpha>1$ atunci $M(X)=\dfrac{\alpha}{\alpha-1}c$, iar dacă $\alpha>2$ atunci:

$$D(X)=\frac{\alpha}{(\alpha-1)^2(\alpha-2)}c^2$$

Dacă $\alpha\in(0,1]$ atunci variabila aleatoare X nu are medie, iar dacă $\alpha\in(0,2]$ variabila aleatoare X nu are dispersie.

5. (*Repartiţia Laplace*). O variabilă aleatoare X se spune că are o repartiţie Laplace, sau repartiţie exponenţială bilaterală, dacă această variabilă aleatoare este de tip continuu şi are funcţia de densitate de forma:

$$f(x)=\frac{\lambda}{2}e^{-\lambda|x|}\ ;\ (\forall)x\in\mathbb{R}$$

unde λ este un număr real pozitiv.

Să se determine funcţia de repartiţie media şi dispersia unei variabile aleatoare cu repartiţie Laplace.

Răspuns:

$$M(X)=0\ ,\ D(X)=\frac{2}{\lambda^2}$$

$$F(x)=\begin{cases}1-\dfrac{1}{2}e^{-\lambda x}\ ;\ x>0\\[2ex]\dfrac{1}{2}e^{\lambda x}\qquad\ ;\ x\le 0\end{cases}$$

6. (*Repartiţia Cauchy*). Spunem că o variabilă aleatoare are o repartiţie Cauchy dacă este de tip continuu şi are funcţia de densitate de forma:

$$f(x)=\frac{1}{\pi}\cdot\frac{c}{c^2+(x-a)^2}\ ,\ (\forall)x\in\mathbb{R}$$

unde a este un număr real oarecare, iar $c>0$. Dacă X are o repartiţie Cauchy

să se arate că variabila aleatoare $y = \frac{X-a}{c}$ are deasemenea o repartiţie Cauchy în care parametrii a şi c au valori 0 şi 1, respectiv.

Să se arate că o variabilă aleatoare cu repartiţie Cauchy nu are moment iniţial de ordin k, oricare ar fi $k \geq 1$.

7. Dacă X este o variabilă aleatoare care are funcţia de repartiţie F continuă şi strict crescătoare, să se arate că variabila aleatoare $Y = F(X)$ are repartiţie uniformă pe intervalul $[0,1]$.

8. Să se arate că, pentru o variabilă aleatoare cu repartiţie lognormală de parametri a şi σ, funcţia de repartiţie se poate calcula după formula:

$$F(x) = \frac{1}{2} + \Phi\left(\frac{\ln x - \ln a}{\sigma}\right)$$

unde Φ este funcţia lui Laplace.

9. (*Repartiţia Gamma*). Spunem că o variabilă aleatoare X are o repartiţie Gamma de parametri $a, b\,(a>o,\ b>0)$, sau că are o repartiţie $\gamma(a,b)$, dacă funcţia sa de densitate este dată prin:

$$f(x) = \begin{cases} 0\ ; & x<0 \\ \dfrac{b^a}{\Gamma(a)} x^{a-1} e^{-bx}\ ; & x \geq 0 \end{cases}$$

unde Γ este funcţia Gamma a lui Euler.

Să se arate că funcţia f este o funcţie de densitate; să se calculeze media şi dispersia unei variabile aleatoare cu repartiţie $\gamma(a,b)$

Răspuns: $M(X) = \frac{a}{b}$, $D(X) = \frac{a}{b^2}$

10. (*Repartiţia Beta*). Se spune că o variabilă aleatoare X are repartiţie Beta de parametri $a, b\,(a>0,\ b>0)$ dacă funcţia sa de densitate este dată prin:

$$f(x) = \begin{cases} 0 & ;\ x \notin (0,1) \\ \dfrac{1}{\beta(a,b)} x^{a-1}(1-x)^{b-1} & ;\ x \in (0,1) \end{cases}$$

Să se arate că f este o funcţie de densitate şi să se calculeze media şi dispersia variabilei aleatoare corespunzătoare.

Răspuns: $M(X) = \frac{a}{a+b}$, $D(X) = \frac{ab}{(a+b)^2(a+b+1)}$

§ 9. LEGEA NUMERELOR MARI

În acest paragraf vom enunţa o serie de rezultate prin care se evidenţiază unele proprietăţi ale şirurilor de variabile aleatore, deosebit de utile în statistica matematică. Înainte de a prezenta astfel de rezultate, vom considera un exemplu în urma căruia va apare evident un rezultat de o semnificaţie deosebită pentru statistica matematică, şi anume teorema lui Bernoulli.

Să presupunem că într-o urnă sînt bile de două culori, de exemplu albe şi negre, necunoscîndu-se proporţia de bile albe existente în această urnă. Dacă notăm cu p probabilitatea de a extrage o bilă albă din acestă urnă, atunci este clar că $(100\ p)\%$ din bilele din urnă sînt albe, restul fiind negre. Prin urmare, dacă am dori să cunoaştem proporţia de bile albe din această urnă ar fi suficient să cunoaştem probabilitatea de extragere a unei bile albe din urnă.

De regulă, pentru determinarea unei valori aproximative a proporţiei de bile albe din urnă se procedează în felul următor: se extrag din urnă un număr oarecare de bile, să spunem n bile, cu sau fără revenire, aproximîndu-se valoarea reală a lui p prin raportul dintre numărul de bile albe extrase şi numărul total de bile extrase. Altfel spus, valoarea reală a lui p este aproximată prin frecvenţa relativă a numărului de bile albe obţinute în cele n extrageri.

Din punct de vedere practic este deosebit de important să ştim ce eroare se comite dacă pentru aproximarea valorii reale a lui p se procedează în acest mod. Dacă pentru moment considerăm că numărul de extrageri, n, este fixat, atunci numărul de bile obţinute în urma celor n extrageri, (frecvenţa absolută a bilelor albe), este o variabilă aleatoare discretă, X_n, ale cărui valori reprezintă numărul de bile albe care se pot obţine în urma celor n extrageri. După cum am mai văzut, pentru n fixat, variabila aleatoare X_n are o repartiţie binomială, în cazul cînd extragerile se fac cu revenire, şi o repartiţie hipergeometrică, în cazul cînd extragerile se fac fără revenire.

Deoarece frecvenţa absolută a numărului de bile albe în n extrageri din urnă este o variabilă aleatoare, urmează că şi frecvenţa relativă a numărului de bile albe este de asemenea o variabilă aleatoare, aceasta fiind variabila aleatoare

$\frac{1}{n} X_n$.

În concluzie, eroarea comisă atunci cînd se aproximează p cu frecvenţa relativă a numărului de bile albe în n extrageri este aleatoare, valoarea acestei erori fiind oricare dintre valorile variabilei aleatoare $\left|\frac{1}{n} X_n - p\right|$. Dacă se fixează un număr $\varepsilon > 0$, atunci probabilitatea ca eroarea absolută să nu depăşească ε este:

$$P\left(\left|\frac{1}{n} X_n - p\right| < \varepsilon\right)$$

şi ţinînd cont de inegalitatea lui Cebîşev, putem scrie:

$$P\left(\left|\frac{1}{n} X_n - p\right| < \varepsilon\right) \geq 1 - \frac{D\left(\frac{1}{n} X_n\right)}{\varepsilon^2}$$

În cazul cînd extragerile se fac cu revenire, se ştie că $D(X_n) = npq$, iar în cazul cînd extragerile se fac fără revenire, $D(X_n) = npq\,\frac{t-n}{t-1}$, unde $q = 1 - q$ iar t este numărul total de bile din urnă. Deci:

$$(*) \quad P\left(\left|\frac{1}{n} X_n - p\right| < \varepsilon\right) \geq 1 - \frac{pq}{n\varepsilon^2}$$

dacă cele n extrageri se fac cu revenire, şi în cazul cînd extragerile se fac cu revenire, t fiind numărul total de bile din urnă:

$$(**) \quad P\left(\left|\frac{1}{n} X_n - p\right| < \varepsilon\right) \geq 1 - \frac{pq}{n\varepsilon^2} \cdot \frac{t-n}{t-1}$$

Din relaţiile (*), (**) se observă că atunci cînd numărul de extrageri este suficient de mare, probabilitatea ca frecvenţa relativă a numărului de bile albe să se abată cu cel mult ε, $\varepsilon > 0$, de la p tinde să se apropie din ce în ce mai mult de 1.

Din aceleaşi relaţii (*), (**) observăm că este posibil să determinăm n astfel încît această eroare absolută să nu depăşească o valoare dată ε, cu o

probabilitate minimă dată $1-\delta$, oricît de apropiată de 1.

Într-adevăr, deoarece $p, q \in [0,1]$ şi $p+q=1$ se constată imediat $0 \le pq \le \frac{1}{4}$ şi atunci, în cazul extragerilor cu revenire, n ar putea fi ales astfel încît:

$$1-\frac{1}{4n\varepsilon^2} \ge 1-\delta$$

iar în cazul cînd extragerile se fac fără revenire, n trebuie ales astfel încît

$$1-\frac{1}{4n\varepsilon^2}\cdot\frac{t-n}{t-1} \ge \delta$$

Exemplul 7.9.1.

Fie o urnă în care se află bile albe şi bile negre. Cîte bile ar trebui să se extragă din urnă, după schema bilei revenite, astfel încît, cu probabilitatea cel puţin 0,5, frecvenţa relativă a numărului de bile extrase să se abată de la probabilitatea de extragere a unei bile albe cu cel mult 0,05.

Pe baza celor de mai sus rezultă că numărul de extrageri n trebuie astfel ales încît:

$$1-\frac{1}{4n(0,05)^2} \ge 0,5$$

de unde urmează $n \ge 100$.

Deci, dacă extragem din urnă cel puţin 100 bile, frecvenţa relativă a bilelor albe se va abate de la probabilitatea p, de a extrage o bilă albă din urnă, cu cel mult 0,05, cu probabilitatea cel puţin 0,5.

Din relaţiile (*), (**) se observă că atunci cînd $n \to \infty$, respectiv $n \to t$, probabilitatea unei abateri oricît de mici a frecvenţei relative de la probabilitatea p tinde să se apropie de 1.

Putem spune atunci că şirul de variabile aleatoare $\left(\frac{1}{n}X_n\right)_{n\in\mathbb{N}}$ tinde la p cu o probabilitate tinzînd la 1.

Aceasta conduce la următoarea definiţie:

Definiţia 7.9.1.

Fie (X_n) un şir de variabile aleatoare definite peste un cîmp de probabilitate (Ω, K, P). Vom spune că şirul (X_n) converge în probabilitate la variabila aleatoare X, definită peste acelaşi cîmp de probabilitate, dacă:

$$\lim_{n\to\infty} P(|X_n - X| < \epsilon) = 1 \; ; \; (\forall)\, \varepsilon > 0 \qquad 7.9.1$$

Definiţia 7.9.2.

Vom spune că un şir de variabile aleatoare (X_n) converge în relaţie la o variabilă aleatoare X dacă:

$$\lim_{n\to\infty} F_n(x) = F(x) \; ; \; (\forall) x \in \mathbb{R} \qquad 7.9.2$$

unde $F(x)$ este funcţie de repartiţie a variabilei X_n iar F este funcţia de repartiţie a variabilei aleatoare X.

Definiţia 7.9.3.

Se spune că un şir de variabile aleatoare (X_n), definite peste un cîmp de probabilitate (Ω, K, P), este supus legii numerelor mari dacă există un şir de constante (a_n) astfel încît:

$$\lim_{n\to\infty} P(|X_n - a_n| < \varepsilon) = 1 \; ; \; (\forall)\, \varepsilon > 0 \qquad 7.9.3$$

Cu alte cuvinte, un şir de variabile aleatoare (X_n) este supus legii numerelor mari pentru n suficient de mare variabila aleatoare X_n poate fi considerată ca fiind o constantă, probabilitatea ca această variabilă aleatoare să se abată de la constanta respectivă cu un număr oricît de mic fiind foarte apropiată de 1.

Teorema lui Bernoulli

Fie A un eveniment oarecare şi fie $f_n(A)$ frecvenţa relativă a evenimentului A în n probe independente. Atunci are loc:

BCU CLUJ CENTRALA UNIVERSITARA LIBRARY

$$\lim_{n\to\infty} P(|f(A) - P(A)| < \varepsilon) = 1 \; ; \; (\forall)\, \varepsilon > 0$$

unde $P(A)$ este probabilitatea evenimentului A.

Demonstraţie: Dacă notăm $P(A) = p$ iar $f_n(A) = \frac{1}{n} X_n$ unde X_n este variabila aleatoare care ia valori, frecvenţele absolute ale producerii evenimentului A în cele n probe independente, atunci X_n are o repartiţie binomială de parametru p. Din cele arătate mai sus rezultă:

$$P(|f_n(A) - p| < \varepsilon) \geq 1 - \frac{pq}{n\varepsilon^2} \; ; \; (\forall)\, \varepsilon > 0 \qquad 7.9.4$$

Trecînd la limită în acestă relaţie se obţine:

$$\lim_{n\to\infty} P(|f_n(A) - P(A)| < \varepsilon) = 1 \; ; \; (\forall)\, \varepsilon > 0 \qquad 7.9.5$$

ceea ce demonstrează teorema.

Să observăm că prin această teoremă se arată că şirul de variabile aleatoare $(f_n(A))$ este supus legii numerelor mari, şirul de constante (a_n) fiind determinat prin $a_n = P(A)$, $n = 1, 2, \ldots$.

Teorema lui Bernoulli este un caz particular pentru o teoremă mai generală, pe care o enunţăm în continuare.

Teorema lui Cebîşev

Fie (X_n) un şir de variabile aleatoare independente cu următoarele proprietăţi:

1. $M(X_n) = m \; ; \; (\forall)\, n$
2. $D(X_n) \leq C \; ; \; (\forall n$

În aceste condiţii şirul de variabile aleatoare (Y_n), $Y_n = \frac{1}{n} \sum_{i=1}^{n} X_i$, converge în probabilitate la m.

Demonstraţie: Din definiţia şirului (Y_n) urmează:

$$M(Y_n) = m$$

şi deoarece, variabilele aleatoare X_i sînt independente putem scrie:

$$D(Y_n) = \frac{1}{n^2} \sum_{i=n}^{n} D(X_i) \leq \frac{C}{n}$$

Folosind inegalitatea lui Cebîşev urmează atunci:

$$P(|Y_n - m| < \varepsilon) \geq 1 - \frac{C}{n\varepsilon^2} \; ; \; (\forall)\, \varepsilon > 0$$

de unde:

$$\lim_{n \to \infty} P(|Y_n - m| < \varepsilon) = 1 \; ; \; (\forall)\, \varepsilon > 0$$

ceea ce demonstrează teorema.

În încheiere vom enunţa o teoremă deosebit de importantă pentru aplicaţiile practice ale statisticii matematice, numită *teorema limită centrală a lui Leapunov*.

Teoremă

Fie (X_n) un şir de variabile aleatoare pentru care sînt îndeplinite condiţiile:

1. Pentru orice n, variabilele aleatoare $X_1, \ldots, X_n$ sînt independente.
2. Pentru orice k, există $M(X_k)$, $D(X_k)$, $M(|X_k - M(X_k)_3|)$
3. Dacă notăm $\sigma_k^2 = D(X_k)$, $H_k(|X_k - M(X_k)|^3)$ atunci:

$$\lim_{n \to \infty} \frac{\left(\sum_{k=1}^{n} H_k \right)^{\frac{1}{3}}}{\left(\sum_{k=1}^{n} \sigma^{2}{}_{k} \right)^{\frac{1}{2}}} = 0$$

BCU IAŞI/CENTRAL UNIVERSITY LIBRARY

În aceste condiții, notăm cu $Y_n = \sum_{k=1}^{n} X_n$, șirul de variabile aleatoare (X_n) converge în repartiție la o variabilă aleatoare cu repartiție normală.

Observăm că între ipotezele acestei teoreme nu este inclusă nici o condiție relativ la repartiția variabilelor aleatoare $X_k, k = 1, 2, \ldots$. În partea de statistică matematică vom vedea cît de utilă este acestă teoremă pentru aplicații practice și, totodată,rolul deosebit pe care îl are repartiția normală între toate repartițiile de probabilitate.

BCU IAȘI/CENTRAL UNIVERSITY LIBRARY

Tabelul 1 Valorile funcţiei lui Laplace:

$$\Phi(z) = \frac{1}{\sqrt{2\pi}} \int_0^z e^{-\frac{x^2}{2}} dx$$

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.0000	0.0040	0.0080	0.0120	0.0160	0.0199	0.0239	0.0279	0.0319	0.0359
0.1	0.0398	0.0438	0.0478	0.0517	0.0557	0.0596	0.0636	0.0675	0.0714	0.0753
0.2	0.0793	0.0832	0.0871	0.0910	0.0948	0.0987	0.1026	0.1064	0.1103	0.1141
0.3	0.1179	0.1217	0.1255	0.1293	0.1331	0.1368	0.1406	0.1443	0.1480	0.1517
0.4	0.1554	0.1591	0.1628	0.1664	0.1700	0.1736	0.1772	0.1808	0.1844	0.1879
0.5	0.1915	0.1950	0.1985	0.2019	0.2057	0.2088	0.2123	0.2157	0.2190	0.2224
0.6	0.2257	0.2291	0.2324	0.2357	0.2389	0.2422	0.2454	0.2486	0.2518	0.2549
0.7	0.2580	0.2612	0.2642	0.2673	0.2704	0.2734	0.2764	0.2794	0.2823	0.2852
0.8	0.2881	0.2910	0.2939	0.2967	0.2995	0.3023	0.3051	0.3078	0.3106	0.3133
0.9	0.3159	0.3186	0.3212	0.3238	0.3264	0.3289	0.3315	0.3340	0.3365	0.3389
1.0	0.3413	0.3438	0.3461	0.3485	0.3508	0.3531	0.3554	0.3577	0.3599	0.3621
1.1	0.3643	0.3665	0.3686	0.3708	0.3729	0.3749	0.3770	0.3790	0.3810	0.3830
1.2	0.3849	0.3869	0.3888	0.3907	0.3925	0.3944	0.3962	0.3980	0.3997	0.4015

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
1.3	0.4032	0.4049	0.4066	0.4082	0.4099	0.4115	0.4131	0.4147	0.4162	0.4177
1.4	0.4192	0.4207	0.4222	0.4236	0.4251	0.4265	0.4279	0.4292	0.4306	0.4319
1.5	0.4332	0.4345	0.4357	0.4370	0.4382	0.4394	0.4406	0.4418	0.4429	0.4441
1.6	0.4452	0.4463	0.4474	0.4484	0.4495	0.4505	0.4515	0.4525	0.4535	0.4545
1.7	0.4554	0.4564	0.4573	0.4582	0.4591	0.4599	0.4608	0.4616	0.4625	0.4633
1.8	0.4641	0.4649	0.4656	0.4664	0.4671	0.4678	0.4686	0.4693	0.4699	0.4706
1.9	0.4713	0.4726	0.4726	0.4732	0.4738	0.4744	0.4750	0.4756	0.4761	0.4767
2.0	0.4772	0.4778	0.4783	0.4788	0.4793	0.4798	0.4803	0.4808	0.4812	0.4817
2.1	0.4821	0.4826	0.4830	0.4834	0.4838	0.4842	0.4846	0.4850	0.4854	0.4857
2.2	0.4861	0.4864	0.4868	0.4841	0.4875	0.4878	0.4881	0.4884	0.4887	0.4890
2.3	0.4893	0.4896	0.4898	0.4901	0.4904	0.4906	0.4909	0.4911	0.4913	0.4916
2.4	0.4918	0.4920	0.4922	0.4925	0.4927	0.4929	0.4931	0.4932	0.4934	0.4936
2.5	0.4938	0.4940	0.4941	0.4943	0.4945	0.4946	0.4948	0.4949	0.4951	0.4952
2.6	0.4953	0.4955	0.4956	0.4957	0.4959	0.4960	0.4961	0.4962	0.4963	0.4964
2.7	0.4965	0.4966	0.4967	0.4968	0.4969	0.4970	0.4971	0.4972	0.4973	0.4974
2.8	0.4974	0.4975	0.4976	0.4977	0.4977	0.4978	0.4979	0.4979	0.4980	0.4981

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
2.9	0.4981	0.4982	0.4982	0.4983	0.4984	0.4984	0.4985	0.4985	0.4986	0.4986
3.0	0.49865	0.4987	0.4987	0.4988	0.4998	0.4989	0.4989	0.4989	0.4990	0.4990
4.0	0.49997									

Exemplu: Dacă $z = 1,16$ atunci $\Phi(z) = 0,3770$.

BCU CLUJ CENTRAL UNIVERSITY LIBRARY

Tabelul 2 Valorile lui ε în funcţie de probabilitatea $\delta = P(X > \varepsilon)$, unde X este o variabilă aleatoare cu repartiţie hi-pătrat şi n grade de libertate, $(n \leq 30)$.

n/δ	0.20	0.10	0.05	0.02	0.01
1	1.642	2.706	3.841	5.412	6.635
2	3.219	4.605	5.991	7.824	9.210
3	4.642	6.251	7.815	9.837	11.345
4	5.989	7.779	9.488	11.668	13.277
5	7.289	9.236	11.070	13.388	15.086
6	8.558	10.645	12.592	15.033	16.812
7	9.803	12.017	14.067	16.622	18.475
8	11.030	13.362	15.507	18.168	20.090
9	12.242	14.684	16.919	19.679	21.666
10	13.442	15.987	18.307	21.161	23.209
11	14.631	17.275	19.675	22.618	24.725
12	15.812	18.549	21.026	24.054	26.217
13	16.985	19.812	22.362	25.472	27.688

BCU IAŞI CENTRALA UNIVERSITARĂ LIBRARY

n/δ	0.20	0.10	0.05	0.02	0.01
14	18.151	21.064	23.685	26.873	29.141
15	19.311	22.307	24.996	28.259	30.578
16	20.465	23.542	26.296	29.633	32.000
17	21.615	24.769	27.587	30.995	33.409
18	22.760	25.989	28.869	32.346	34.805
19	23.900	27.204	30.144	33.687	36.191
20	25.038	28.412	31.410	35.090	37.566
21	26.171	29.615	32.671	36.343	38.932
22	27.301	30.813	33.924	37.659	40.289
23	28.429	32.007	35.172	38.968	41.638
24	29.553	33.196	36.415	40.270	42.980
25	30.675	34.382	37.652	41.566	44.314
26	31.795	35.563	38.885	42.856	45.642
27	32.912	36.741	40.113	44.140	46.963
28	34.027	37.916	40.337	45.419	48.278

BOGAZICI UNIVERSITY LIBRARY

n/δ	0.20	0.10	0.05	0.02	0.01
29	35.139	39.087	42.557	46.963	49.588
30	26.150	40.256	43.773	47.962	50.892

Exemplu: Dacă $\delta = 0,05$ şi $n = 5$ atunci $\varepsilon = 11,070$.

BCU "LUCIAN BLAGA" CENTRAL UNIVERSITY LIBRARY

Tabelul 3 Valorile lui ε în funcţie de probabilitatea $\delta = P(|X| > \varepsilon)$, unde X este o variabilă aleatoare cu repartiţie t şi n grade de libertate, $(n \leq 30)$.

n/δ	0.10	0.05	0.02	0.01
1	6.3114	12.706	31.821	63.657
2	2.920	4.303	6.965	9.925
3	2.353	3.182	4.541	5.841
4	2.132	2.776	3.747	4.604
5	2.015	2.571	3.365	4.032
6	1.943	2.447	3.143	3.707
7	1.895	2.365	2.998	3.499
8	1.860	2.306	2.896	3.355
9	1.833	2.262	2.821	3.250
10	1.812	2.228	2.764	3.169
11	1.796	2.201	2.718	3.106
12	1.782	2.179	2.681	3.055
13	1.771	2.160	2.650	3.012

BCU IASI CENTRALA UNIVERSITARA LIBRARY

n/δ	0.10	0.05	0.02	0.01
14	1.761	2.145	2.624	2.977
15	1.753	2.131	2.602	2.947
16	1.746	2.120	2.583	2.921
17	1.740	2.110	2.567	2.898
18	1.734	2.101	2.552	2.878
19	1.729	2.093	2.539	2.861
20	1.725	2.086	2.528	2.845
21	1.721	2.080	2.518	2.831
22	1.717	2.074	2.508	2.819
23	1.714	2.069	2.500	2.807
24	1.711	2.064	2.492	2.797
25	1.708	2.060	2.485	2.787
26	1.706	2.056	2.479	2.779
27	1.703	2.052	2.473	2.771
28	1,701	2.048	2.467	2.763

BCU ASICENTRAL UNIVERSITY LIBRARY

n/δ	0.10	0.05	0.02	0.01
29	1.699	2.045	2.462	2.756
30	1.697	2.042	2.457	2.750

Exemplu: Dacă $\delta = 0,05$ şi $n = 10$ atunci $\varepsilon = 2,228$.

BCU IASI CENTRAL UNIVERSITY LIBRARY

BIBLIOGRAFIE

1. E. Boroş, D. Opriş - *Introducere în optimizarea liniară şi aplicaţii*, Ed. Facla, Timişoara 1979.
2. I. Craiu s.a. - *Matematici pentru economişti*, vol. 1, 2, 3. Ed. Tehnică, Bucureşti 1971.
3. I. Creangă, C. Haimovici - *Algebră liniară*, E.S.D.P. Buc. 1962.
4. B. Demidovitch s.a. - *Recueil d'exercices et de problèmes d'analyse mathématique*, Ed. Mir, Moscou 1971.
5. G. Fihtengolţ - *Bazele analizei matematice*, vol. 2,1, Moscova 1955.
6. Ion D. Ion, N. Radu - *Algebră*, E.D.P., Bucureşti 1975.
7. H. Ionescu - *Statistică matematică*, E.D.P., Bucureşti 1962.
8. N. Mihăilă, O. Popescu - *Matematici speciale aplicate în economie*, E.D.P., Bucureşti 1978.
9. C. Reischer, G. Sâmboan, R. Teodorescu - *Teoria probabilităţilor*, E.D.P., Bucureşti 1967.
10. C. Smadici - *Matematici speciale aplicate în economie*, Iaşi 1983.
11. V. Smirnov - *Cours de mathématiques supérieurs*, t.I,II. Ed. Mir, Moscou 1972.
12. V. Tamaş - *Lecţii de calcul economic*, Iaşi 1970.

BCU IASI/CENTRAL UNIVERSITY LIBRARY

635 / Popescu Dumitru
418474

Redactor: Christian TĂMAŞ

Culegerea şi procesarea textului:
A'92 srl, Tel&Fax 098-140270

Tehnoredactarea computerizată:
Radu C. ASANDEI

Apărut: 1993
Format: 54×84/16
Coli tipo: 17
Bun de tipar: 5.10.1993

Editura GRAPHIX
Str. Ştefan cel Mare nr.2 bis IAŞI

TIPARUL EXECUTAT
SUB CDA Nr. 450

©

ROTA SRL
J 22-672-91
Bd. ŞTEFAN CEL MARE ŞI SFÎNT 2 BIS
6600 IAŞI * ROMANIA
Phone: +40-98-116981
Fax: +40-98-117631
COD FISCAL: 1963149
SIRUES: 225556195
DGFP: 277/91

BCU IASI/CENTRAL UNIVERSITY LIBRARY

BCU IASI/CENTRAL UNIVERSITY LIBRARY

BCU IASI/CENTRAL UNIVERSITY LIBRARY